


```
SQ SEQUENCE 583 AA; 67459 MW; 318DC4D81CD0FF2A CRC64;
Query Match 100.0%; Score 3209; DB 4; Length 583;
Best Local Similarity 100.0%; Pred. No. 7,7e-212;
Matches 583; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSYGEIEGKFLGPREEVTSERCKKLKSTTSYVFHNSNADFHRIQKTDGNDWVPVTII 60
DB 1 MSYGEIEGKFLGPREEVTSERCKKLKSTTSYVFHNSNADFHRIQKTDGNDWVPVTII 60
QY 61 DVRGHSYLQENKIKTTDLRPLHDEMPGNDPVDIESIDSVLQEARPLVSADDEIYSTS 120
DB 61 DVRGHSYLQENKIKTTDLRPLHDEMPGNDPVDIESIDSVLQEARPLVSADDEIYSTS 120
QY 121 KAFITGPIYKPEKKRNRGRNEAHVNLGINDRGGQKQKFNSEKSEIDNLFQFYKIE 180
DB 121 KAFITGPIYKPEKKRNRGRNEAHVNLGINDRGGQKQKFNSEKSEIDNLFQFYKIE 180
QY 181 ELEKEKDFGNSCKESPSQOFVFPFYEGHNNGLLKPDDEKKDLNKAAMPSCDYQOQNLG 240
DB 181 ELEKEKDFGNSCKESPSQOFVFPFYEGHNNGLLKPDDEKKDLNKAAMPSCDYQOQNLG 240
QY 241 NEPKYPCNGQVITPFCDTSTSPERPEWQSVYFIVPGPLPSLNYHLNLTQRFSGPPNP 300
DB 241 NEPKYPCNGQVITPFCDTSTSPERPEWQSVYFIVPGPLPSLNYHLNLTQRFSGPPNP 300
QY 301 PSNIFQAODDSQIQNGYVYNNCHVNNWCMTPDQNNXYTDCSENSSVHPSGCGSMQDRY 360
DB 301 PSNIFQAODDSQIQNGYVYNNCHVNNWCMTPDQNNXYTDCSENSSVHPSGCGSMQDRY 360
QY 361 VSNCFCEVREKRCWKDCHMDKHNGTDRFVNOQFOBEKLNKLOKLLILLRGLPGSGKTTLSR 420
DB 361 VSNCFCEVREKRCWKDCHMDKHNGTDRFVNOQFOBEKLNKLOKLLILLRGLPGSGKTTLSR 420
QY 421 ILLGNRQDGI VFTDDYFHQDGYRYNNQGLGDAHDNQNRAKQAIQDGRSPVIIDNTNI 480
DB 421 ILLGNRQDGI VFTDDYFHQDGYRYNNQGLGDAHDNQNRAKQAIQDGRSPVIIDNTNI 480
QY 481 QAWEMKPYVEVAIGKYRVEHEPETWKKFDPPELEKRNKHGVSRRKKTQAQMLDRYEYQMS 540
DB 481 QAWEMKPYVEVAIGKYRVEHEPETWKKFDPPELEKRNKHGVSRRKKTQAQMLDRYEYQMS 540
QY 541 ISIVMNSVEPSHKSTQRPPTPPQGRQWRGSGSLGSHNRVCVTNNH 583
DB 541 ISIVMNSVEPSHKSTQRPPTPPQGRQWRGSGSLGSHNRVCVTNNH 583

RESULT 2
Q91ZG4 PRELIMINARY; PRT; 575 AA.
AC Q91ZG4; 2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE D41766.4 (novel protein).
GN D41766.4
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Wall M.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL358892; CAC42205.1;
SQ SEQUENCE 575 AA; 66250 MW; C41FAA8CC3EE4408 CRC64;

Query Match 61.5%; Score 1975; DB 11; Length 575;
Best Local Similarity 65.2%; Pred. No. 2,7e-127;
Matches 380; Conservative 70; Mismatches 125; Indels 8; Gaps 7;

QY 1 MSYGEIEGKFLGPREEVTSERCKKLKSTTSYVFHNSNADFHRIQKTDGNDWVPVTII 60
DB 1 MSYGEIEGKFLGPREEVTSERCKKLKSTTSYVFHNSNADFHRIQKTDGNDWVPVTII 60
```

```
DB 1 MPYSEVEAKFLGPKGEQTRPCYKKLKSAADDGVSPLRGDPDIHRIQKPRNNRVAVATI 60
QY 61 DVRGHSYLQENKIKTTDLRPLHDEMPGNDPVDIESIDSVLQEARPLVSADDEIYSTS 120
DB 61 DVRGHSYLQENKIKTTDLRPLHDEMPGNDPVDIESIDSVLQEARPLVSADDEIYSTS 120
QY 121 KAFITGPIYKPEKKRNRGRNEAHVNLGINDRGGQKQKFNSEKSEIDNLFQFYKIE 180
DB 121 KAFITGPIYKPEKKRNRGRNEAHVNLGINDRGGQKQKFNSEKSEIDNLFQFYKIE 180
QY 181 ELEKEKDFGNSCKESPSQOFVFPFYEGHNNGLLKPDDEKKDLNKAAMPSCDYQOQNLG 240
DB 181 ELEKEKDFGNSCKESPSQOFVFPFYEGHNNGLLKPDDEKKDLNKAAMPSCDYQOQNLG 240
QY 241 NEPKYPCNGQVITPFCDTSTSPERPEWQSVYFIVPGPLPSLNYHLNLTQRFSGPPNP 300
DB 241 NEPKYPCNGQVITPFCDTSTSPERPEWQSVYFIVPGPLPSLNYHLNLTQRFSGPPNP 300
QY 301 PSNIFQAODDSQIQNGYVYNNCHVNNWCMTPDQNNXYTDCSENSSVHPSGCGSMQDRY 360
DB 301 PSNIFQAODDSQIQNGYVYNNCHVNNWCMTPDQNNXYTDCSENSSVHPSGCGSMQDRY 360
QY 361 VSNCFCEVREKRCWKDCHMDKHNGTDRFVNOQFOBEKLNKLOKLLILLRGLPGSGKTTLSR 420
DB 361 VSNCFCEVREKRCWKDCHMDKHNGTDRFVNOQFOBEKLNKLOKLLILLRGLPGSGKTTLSR 420
QY 421 ILLGNRQDGI VFTDDYFHQDGYRYNNQGLGDAHDNQNRAKQAIQDGRSPVIIDNTNI 480
DB 421 ILLGNRQDGI VFTDDYFHQDGYRYNNQGLGDAHDNQNRAKQAIQDGRSPVIIDNTNI 480
QY 481 QAWEMKPYVEVAIGKYRVEHEPETWKKFDPPELEKRNKHGVSRRKKTQAQMLDRYEYQMS 540
DB 481 QAWEMKPYVEVAIGKYRVEHEPETWKKFDPPELEKRNKHGVSRRKKTQAQMLDRYEYQMS 540
QY 541 ISIVMNSVEPSHKSTQRPPTPPQGRQWRGSGSLGSHNRVCVTNNH 583
DB 541 ISIVMNSVEPSHKSTQRPPTPPQGRQWRGSGSLGSHNRVCVTNNH 583

RESULT 3
Q9NVK2 PRELIMINARY; PRT; 659 AA.
AC Q9NVK2; 2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CDNA FLJ10680 f1s, clone NT2RP2006573, weakly similar to 2',3'-cyclic
DE nucleotide 3'-phosphodiesterase (EC 3.1.4.37) (Fragment).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoka S., Ishii S., Kawai Y.,
RA Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K.,
RA Masubo Y., Kanehori K.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK001542; BAA91748.1;
DR InterPro; IPR001899; Gram_pos_anchor.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
FT NON_TER 659
SQ SEQUENCE 659 AA; 73887 MW; 9974015C7D6EAC1F CRC64;

Query Match 15.9%; Score 509; DB 4; Length 659;
Best Local Similarity 27.9%; Pred. No. 9,2e-27;
Matches 165; Conservative 88; Mismatches 157; Indels 182; Gaps 28;

QY 45 RIQKTDGNDWVPVTIIIVRGHSYLQENKIKTTD---LHRPLHDEMPGNDPVTII 95
DB 45 RIQKTDGNDWVPVTIIIVRGHSYLQENKIKTTD---LHRPLHDEMPGNDPVTII 95
```

```

Db 93 KIESSSOSFV-----ASEN0YGAESKIMEKREPESEDSKMSDFLMOQLTE 140
Qy 96 SIDSOVLDEARPLVSA--DEEITS-----TSKAFICPIYKPKKKRNGRNEA 143
Db 141 DIDS-LIQNAFEKLNSSPDQVSEFLPSQDVNSFNSSEFINP-----SSNMT 188
Qy 144 HV-----LNCIN-DRGQ-----KEKQKFNSEKSI--DNELF--QFYKEIEEL--- 182
Db 189 PIFSTQNMNLNGENLNSSTLSLNLPLPSHVSVLNESKCFIKNTALLESNPEDSLSS 248
Qy 183 -----EKEKGFENSCKE-----SEPOEQFVPFEGHNNGLL 215
Db 249 LNVASDSIAGCSLNGKQKELLESEVEAQFSEAPVDLDASPPQACLNLP-----GLD 301
Qy 216 KP---DEEKKDLNKKAMPSCDYOQNLGNEDPKYPCNGQVIPT-----FCDTSFSTR 265
Db 302 LPTGGDQKSTRVSDVFLPS-----EGFNFKPKHKP---ELPTGKGDVSYCPV----- 346
Qy 266 PEMQSYPIPIVYGPPLPSLNHLNIGRSGPPNPSPNIFQADDSQIONGY--VNNCH 323
Db 347 ---LAPLPLPPPPPPPMWN-----PMIPAFDLFGQ-----NHGFVAPVVTGA 387
Qy 324 VMNMGTFQNNNEYTPDCSENRSVHPSGNGCSMODRYVNGCFEVRCKMDKCMKNG 383
Db 388 AHMRSY-----NYTPPPSVISHTSPT-----KYMRN-----KDG 416
Qy 384 TDRFVNOQ---FO--EEKLNKLQKLLILRLGLPGSGKTLIRLLGQNRDGIYSTDYF 438
Db 417 TSAIYQVETPRVQVYARKTSYVGLVYLLRGLPGSGKSLARLQGDNPSTGILSTDYF 476
Qy 439 HHODGRIYVNOIGDAHDMNONAKAIDOGSPVIIDNTINQAMKRYVEVALKGR 498
Db 477 YINGQYQFVKKYLGEAHENNONAKAEAFKKISPIIIDNTINQAMKRYVVALSQHKYK 536
Qy 499 VERHEPETMMKDPPELEKRNKGVSRRKIAQMLDREYOMSTIYVNSVSP 550
Db 537 VLEFPDITMMKRPKELARRNIHGVSKKITRLEHYQRFVSPITMSSSV 588

```

RESULT 4

```

Q9P2D4 PRELIMINARY: PRT: 1399 AA.
AC Q9P2D4: 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE KIAA1413 protein (Fragment).
GN KIAA1413.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K., Hirotsawa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Ref. 7:65-73(2000).
DR EMBL: AB037834; BAA92651.1; -;
DR InterPro: IPR002625; Smr/MutS2.
DR Pfam: PF01713; Smr: 1.
DR SMART: SM00463; Smr: 1.
FT NON_TER 1
SQ SEQUENCE 1399 AA: 158124 MW: 10ADC0B597C22BE0 CRC64;

```

Query Match 14.6%; Score 469; DB 4; Length 1399;
 Best Local Similarity 47.6%; Pred. No. 1.3e-23;
 Matches 88; Conservative 40; Mismatches 47; Indels 10; Gaps 2;

Qy 371 RCMKHCHMDKHNCTDRFVNO-----OFEBKLNKLQKLLILRLGLPGSGKTLIRLLGLG 425

```

Db 55 KYMRN-----KQGTSAIYQVETPRVQVYARKTSYVGLVYLLRGLPGSGKSLARLQGD 109
Qy 426 NEDGIVFSTDYFHHODGRIYVNOIGDAHDMNONAKAIDOGSPVIIDNTINQAMK 485
Db 110 NPSGVLSTDDYFYINGQYQFVKKYLGEAHENNONAKAEAFKKISPIIIDNTINQAM 169
Qy 466 KRYVEAIVCKGYRVEHEPETMMKDPPELEKRNKGVSRRKIAQMLDREYOMSTIYV 545
Db 170 KRYVALSQKHKKYKLVLEFPDITMMKRPKELARRNIHGVSKKITRLEHYQRFVSP 229
Qy 546 NSVEP 550
Db 230 SSSVP 234

```

RESULT 5

```

O43712 PRELIMINARY: PRT: 87 AA.
AC O43712: 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE X57 protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Mcmurray A., Hunt A.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: 275887; CAB00067.1; -;
FT NON_TER 87
SQ SEQUENCE 87 AA: 10074 MW: 2E350DF672BA3CBA CRC64;

```

Query Match 14.0%; Score 449.5; DB 4; Length 87;
 Best Local Similarity 82.1%; Pred. No. 8.5e-24;
 Matches 87; Conservative 0; Mismatches 0; Indels 19; Gaps 1;

```

Qy 1 MSYGEIEGKFLGPREEVTSEPRCKKLTSTESYVFHNSNADFHRIQETGNDWVPVIT 60
Db 1 MSYGEIEGKFLGPREE-----HNHNSNADFHRIQETGNDWVPVIT 41
Qy 61 DVRGHSYLDENKIKTTDLRPLHDEMPGRPDVIESIDSOVLQEAR 106
Db 42 DVRGHSYLDENKIKTTDLRPLHDEMPGRPDVIESIDSOVLQEAR 87

```

RESULT 6

```

Q91ZG6 PRELIMINARY: PRT: 238 AA.
AC Q91ZG6: 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE DM417G6.2.1 (novel protein (isoform 1)).
GN DM417G6.2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Wall M.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL358892; CAC42202.1; -;
SQ SEQUENCE 238 AA: 28038 MW: 8E55135F6EC4E9F9 CRC64;

```

Query Match 13.3%; Score 427; DB 11; Length 238;
 Best Local Similarity 47.1%; Pred. No. 1.1e-21;
 Matches 82; Conservative 34; Mismatches 54; Indels 4; Gaps 2;

RESULT 9

Q8WTU5 PRELIMINARY; PRT; 173 AA.
 AC Q8WTU5;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical 20.7 kDa protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRIN;
 RA Strausberg R.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC022188; ANH22188.1;
 KW Hypothetical protein.
 SQ SEQUENCE 173 AA; 20678 MW; E79510AF67362A1E CRC64;

Query Match 8.5%; Score 272.5; DB 4; Length 173;
 Best Local Similarity 56.5%; Pred. No. 2.7e-11;
 Matches 52; Conservative 16; Mismatches 23; Indels 1; Gaps 1;

QY 398 NKLOKLLILRLGPGSGKTLRLILGQNRDGIYFTDDYFHHODG-YRYNVNQLGDADH 456

DB 38 HSFRRKLYLLRGLPGSGKTLRLILOHDFPRALIFSTDFEFREDGAYEFNPDLBAHE 97

QY 457 WNONRAKAOIDQGRSPYIINTNIOAMEMKPY 488

DB 98 WNOFRARAKMRNGISPIIINTNLAHEMKPY 129

RESULT 10

Q9Y273 PRELIMINARY; PRT; 137 AA.

AC Q9Y273;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Hypothetical 16.4 kDa protein (92M18.2.1) (Novel protein).
 GN 92M18.2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Rhodes S.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.

RA Hunt A.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL049788; CAB42446.1;
 DR EMBL: AL049785; CAB42443.1;
 DR EMBL: AL049786; CAB42444.1;
 DR EMBL: 273359; CAC94789.1;
 KW Hypothetical protein.

QY SEQUENCE 137 AA; 16393 MW; BC0939347608B77C CRC64;

Query Match 7.7%; Score 247; DB 4; Length 137;
 Best Local Similarity 42.9%; Pred. No. 1.1e-09;
 Matches 48; Conservative 21; Mismatches 39; Indels 4; Gaps 1;

QY 469 GRSVVIIDNTNIOAMEMKPYEVAIGKGYRVEFEPEPTWKFPDELEKRNKHGVSRRKI 528

DB 4 GISPIIDNTNLAHEMKPYAVMALENNYEVIFREPDTRKPFVNOELARRNHGVSREKI 63

QY 529 AQMLDREYOMSISIVANSVEPS---HKSTORPPPOGRORWGSGLSGSHNR 576
 DB 64 HKMKERYEHVTFHSVLAHEKPSRMNRNODRNNALPSSNNARVMYSYEFPPNR 115

RESULT 11

Q8VE50 PRELIMINARY; PRT; 134 AA.

ID Q8VE50;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 21, Last annotation update)
 DE Similar to hypothetical gene CG018.
 GN A1428195.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC019754; ANH19754.1;
 DR MGI: 2140872; A1428195.
 SQ SEQUENCE 134 AA; 15784 MW; 572C8A1F870CD649 CRC64;

Query Match 7.6%; Score 242.5; DB 11; Length 134;
 Best Local Similarity 44.1%; Pred. No. 2.2e-09;
 Matches 45; Conservative 21; Mismatches 33; Indels 3; Gaps 1;

QY 469 GRSVVIIDNTNIOAMEMKPYEVAIGKGYRVEFEPEPTWKFPDELEKRNKHGVSRRKI 528

DB 4 GISPIIDNTNLAHEMKPYAVMALENNYEVIFREPDTRKPFVNOELARRNHGVSREKI 63

QY 529 AQMLDREYOMSISIVANSVEPSHKT---QRPPEPGRORW 567

DB 64 QRMKERYEHVTFHSVLAHEKPSRMNRNODRNNALPSSNNARVMYSYEFPPNR 105

RESULT 12

Q25201 PRELIMINARY; PRT; 232 AA.

ID Q25201;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical protein Hp0454.
 GN Hp0454.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter.
 NCBI_TaxID=210;

RA [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=26695 / ATCC 700392;
 RX MEDLINE=97394467; PubMed=9252185;
 RA Tomb J.-F., White O., Kervilange A.R., Clayton R.A., Sutton G.G.,
 RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
 RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
 RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
 RA McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,
 RA Berg D.E., Gocayne J.D., Uterback T.R., Peterson J.D., Kelley J.M.,
 RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,
 RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
 RA Venter J.C.;

RT "The complete genome sequence of the gastric pathogen Helicobacter
 pylori."
 RL Nature 388:539-547(1997).
 DR EMBL: AE000560; AAD07520.1;
 DR TIGR: HP0454;

QY SEQUENCE 232 AA; 27447 MW; 398AB5A07CA27668 CRC64;

Query Match 6.1%; Score 195.5; DB 16; Length 232;
 Best Local Similarity 32.0%; Pred. No. 7.4e-06;
 Matches 57; Conservative 33; Mismatches 63; Indels 25; Gaps 8;

```

QY 394 EEKLNKLOKLLILLRCLPGSKTTLRSL-----LGNRDGIVFSTDDYFHHQD--GYR 445
DB 8 QKRMNSKLVINRAIPGGKSTSLIKQIEELAKSLGHSIS--VHSTDYFIQTDEEGIR 65
QY 446 YNV--NOLGDAHDWQNRKAQIDGRSPVITDNTNQAWEMKPYVEVAIGKGYRVEFH 502
DB 66 HYVDKKLNEYHQNQAEKALENRIDIVVCDNTNPNESWQSKPYTDWAREFGYKILLI 125
QY 503 EPTWKKFDPPELEKRNKH--VSR--KKIAQMLDRYEQMSISIVMSNVEPSHSTQ 556
DB 126 D-----FKNRHLETPMDYGDVQACIKRPGIAKHVDYFLERVL--VEPDYVEK 175

RESULT 13
Q19804 PRELIMINARY; PRT; 162 AA.
AC Q19804;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Hypothetical 18.4 kDa protein.
GN F26A1.14.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Fulton L.;
RT "The sequence of C. elegans cosmid F26A1.";
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U27312; AAA68256.1; -
KW Hypothetical protein.
SQ SEQUENCE 162 AA; 18438 MW; FC2D9FC15F985140 CRC64;

Query Match 5.1%; Score 164.5; DB 5; Length 162;
Best Local Similarity 26.2%; Pred. No. 0.00063;
Matches 39; Conservative 35; Mismatches 46; Indels 29; Gaps 4;

QY 394 EEKLNKLOKLL-----ILLRGLPGSKTTLRILLGNRDGIVFSTDDYFHHQDGYR 446
DB 13 DAEISDIQKCLVEGHTILINGVTGSKSTLARELVNHSNGVIVKND-----VSN 64
QY 447 NVNQLGDAHDWQNRKAQIDGRSPVITDNTNQAWEMKPYVEVAIGKGYRVEFHEPET 506
DB 65 NITR-----SVRFIDEDKHLIVVDENQVRSRKKFAELAVNGHVEIFVLEPDT 114
QY 507 WKKFDPPELEKRNKH-----GYSRKKIAQM 531
DB 115 DWRHDAIECKRSEKDEBIGSTESKIMQL 143

RESULT 14
Q9KRM6 PRELIMINARY; PRT; 149 AA.
AC Q9KRM6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)

```

```

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Hypothetical protein VC1610.
GN VC1610.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umavam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Olin H., Dragol I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae.";
RL Nature 406:477-483(2000).
DR EMBL; AE004238; AAF94764.1; -
DR TIGR; VC1610; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 149 AA; 17450 MW; 45D56676A4FF8A9A CRC64;

Query Match 5.0%; Score 161; DB 16; Length 149;
Best Local Similarity 29.5%; Pred. No. 0.00098;
Matches 41; Conservative 26; Mismatches 46; Indels 26; Gaps 6;

QY 404 LILLRGLPGSKTTLRILLGNRDGIVFSTDDYFHHQDGYRNVNOLGDAHDWQNR 462
DB 30 LTLIRGLPGSKSTLAKTL-----SAVHLEADMYVNPQGEYHERPEYLAQAHWCQOQT 84
QY 463 KOAIDOGSPVITDNTNQAWEMKPYVEVAIGKGYRVEF-----HEPETWKKFDPPELEK 517
DB 85 EYWLOQGD--VVSNTFVRHMEMAVYRKLA--RQYRAKLTLVCRE-----RY 129
QY 518 RNKGVSRKKIAQMLDRYE 536
DB 130 QNVHGVDEATVERMQOQW 148

RESULT 15
Q91V42 PRELIMINARY; PRT; 400 AA.
AC Q91V42;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cyclic nucleotide phosphodiesterase 1.
GN CNP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ILS, AND ISS;
RA Ehringer M.A., Thompson J., Conroy O., Xu Y., Yang F., Canniff J.,
RA Beeson M., Gordon L., Bennett B., Johnson T.E., Sikela J.M.;
RT "High-Throughput Sequence Identification of Gene Coding Variants
Within Alcohol-Related QTLs.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE ADENYLATE KINASE FAMILY.
DR EMBL; AF332056; AAK56085.1; -
DR EMBL; AF332055; AAK56084.1; -
DR MGI; MGI:88437; Cnpl.
DR InterPro; IPR001230; Prenyl_site.
DR PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
KW Kinase; Transferase.
SQ SEQUENCE 400 AA; 44654 MW; B974404499DE29BE CRC64;

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 18:04:15 ; Search time 12.2667 Seconds
(without alignments)
55.005 Million cell updates/sec

Title: US-09-924-654-4_COPY_550_565
Perfect score: 92
Sequence: 1 PSHKSTGRPPPGGRQ 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 188354 seqs, 42170167 residues

Total number of hits satisfying chosen parameters: 188354

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: Published Applications_AA*

1: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep.*
2: /cgn2_6/prodata/1/pubppa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodata/1/pubppa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/1/pubppa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubppa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/1/pubppa/US07_PUBCOMB.pep.*
7: /cgn2_6/prodata/1/pubppa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/prodata/1/pubppa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/1/pubppa/US09_NEW_PUB.pep.*
10: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep.*
11: /cgn2_6/prodata/1/pubppa/US10_NEW_PUB.pep.*
12: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep.*
13: /cgn2_6/prodata/1/pubppa/US60_NEW_PUB.pep.*
14: /cgn2_6/prodata/1/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	100.0	583	10 US-09-924-654-4	Sequence 4, Appl1
2	92	100.0	594	10 US-09-925-300-1079	Sequence 1079, Ap
3	49	53.3	901	10 US-09-737-149-33	Sequence 33, Appl
4	49	53.3	998	9 US-10-106-534-2	Sequence 2, Appl1
5	49	53.3	1043	10 US-09-737-149-8	Sequence 8, Appl1
6	48.5	52.7	83	9 US-10-028-072-42	Sequence 42, Appl1
7	48.5	52.7	83	9 US-10-121-049-42	Sequence 42, Appl1
8	48.5	52.7	83	9 US-10-123-904-42	Sequence 42, Appl1
9	48.5	52.7	83	9 US-10-140-470-42	Sequence 42, Appl1
10	48.5	52.7	83	9 US-10-140-470-42	Sequence 42, Appl1
11	48.5	52.7	83	9 US-10-176-918-42	Sequence 42, Appl1
12	48.5	52.7	83	9 US-10-176-921-42	Sequence 42, Appl1
13	48.5	52.7	83	9 US-10-137-865-42	Sequence 42, Appl1
14	48.5	52.7	83	9 US-10-140-474-42	Sequence 42, Appl1
15	48.5	52.7	83	9 US-10-142-431-42	Sequence 42, Appl1
16	48.5	52.7	83	9 US-10-143-114-42	Sequence 42, Appl1
17	48.5	52.7	83	9 US-10-140-002-42	Sequence 42, Appl1
18	47	51.1	196	9 US-09-989-920-224	Sequence 224, Appl
19	47	51.1	345	9 US-10-036-729-4	Sequence 4, Appl1

20	46	50.0	80	10 US-09-864-761-41757	Sequence 41757, A
21	46	50.0	468	10 US-09-925-297-736	Sequence 736, App
22	46	50.0	506	10 US-09-801-368-90	Sequence 90, Appl
23	45	48.9	274	10 US-09-850-887-4	Sequence 4, Appl1
24	45	48.9	322	10 US-09-823-038A-33	Sequence 33, Appl1
25	44.5	48.4	21	10 US-09-873-106B-22	Sequence 22, Appl1
26	44.5	48.4	70	10 US-09-873-106B-11	Sequence 11, Appl1
27	44.5	48.4	351	10 US-09-796-033-6	Sequence 6, Appl1
28	44.5	48.4	351	10 US-09-730-465-6	Sequence 6, Appl1
29	44	47.8	72	10 US-09-864-761-40031	Sequence 40031, A
30	44	47.8	107	9 US-10-178-213-269	Sequence 269, App
31	44	47.8	132	10 US-09-864-761-43644	Sequence 43644, A
32	44	47.8	466	12 US-10-095-492-16	Sequence 16, Appl
33	44	47.8	836	10 US-09-934-323-5	Sequence 5, Appl1
34	44	47.8	1175	10 US-09-771-161A-224	Sequence 224, App
35	44	47.8	1175	10 US-09-771-161A-225	Sequence 225, App
36	44	47.8	1175	10 US-09-771-161A-226	Sequence 226, App
37	43	46.7	117	10 US-09-864-761-34970	Sequence 34970, A
38	43	46.7	390	10 US-09-925-302-625	Sequence 625, App
39	43	46.7	491	9 US-10-217-774-2	Sequence 2, Appl1
40	43	46.7	538	9 US-09-976-740-43	Sequence 43, Appl1
41	43	46.7	538	12 US-10-023-529-43	Sequence 43, Appl1
42	43	46.7	538	12 US-10-023-529-43	Sequence 43, Appl1
43	43	46.7	674	9 US-10-086-464-14	Sequence 14, Appl1
44	43	46.7	1224	9 US-10-217-774-4	Sequence 4, Appl1
45	42	45.7	88	10 US-09-867-550-1482	Sequence 1482, Ap

ALIGNMENTS

RESULT 1
US-09-924-654-4
; Sequence 4, Application US/09924654
; Patent No. US20020146712A1
; GENERAL INFORMATION:
; APPLICANT: Goold, Richard D.
; APPLICANT: Aeridom, Ingrid E.
; APPLICANT: Sellhammer, Jeffrey J.
; APPLICANT: Coleman, Roger
; TITLE OF INVENTION: TUMOR SUPPRESSOR
; FILE REFERENCE: PC-0049 CIP
; CURRENT APPLICATION NUMBER: US/09/924,654
; CURRENT FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020146712A1 496071CD1
US-09-924-654-4
Query Match 100.0%; Score 92; DB 10; Length 583;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PSHKSTGRPPPGGRQ 16
DB 550 PSHKSTGRPPPGGRQ 565
RESULT 2
US-09-925-300-1079
; Sequence 1079, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101

```
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1079
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (430)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1079

Query Match          100.0%; Score 92; DB 10; Length 594;
Best Local Similarity 100.0%; Pred. No. 0.00014;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSHKSTQRPQPGRQ 16
   :: |||||
Db 561 PSHKSTQRPQPGRQ 576

RESULT 3
US-09-737-149-33
; Sequence 33, Application US/09737149
; Patent No. US20020077466A1
; GENERAL INFORMATION:
; APPLICANT: Spaderna, Steven K
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Muralidhara, Padigaru
; APPLICANT: Spytek, Kimberly A.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-620 CIP
; CURRENT APPLICATION NUMBER: US/09/737,149
; CURRENT FILING DATE: 2001-06-15
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: 60/170,564
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,165
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,544
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 60/174,404
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: 60/174,962
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 60/223,929
; PRIOR FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 33
; LENGTH: 901
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-737-149-33

Query Match          53.3%; Score 49; DB 10; Length 901;
Best Local Similarity 66.7%; Pred. No. 62;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 STQRPQPGRQ 16
   :: |||||
Db 401 ASARPPPGQAQ 412

RESULT 4
US-09-737-149-8
; Sequence 8, Application US/09737149
; Patent No. US20020077466A1
; GENERAL INFORMATION:
; APPLICANT: Spaderna, Steven K
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Muralidhara, Padigaru
; APPLICANT: Spytek, Kimberly A.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-620 CIP
; CURRENT APPLICATION NUMBER: US/09/737,149
; CURRENT FILING DATE: 2001-06-15
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: 60/170,564
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,165
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,362
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,544
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 60/174,404
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: 60/174,962
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 60/223,929
; PRIOR FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 8
; LENGTH: 1043
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-737-149-8

Query Match          53.3%; Score 49; DB 10; Length 1043;
Best Local Similarity 66.7%; Pred. No. 72;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 STQRPQPGRQ 16
   :: |||||
Db 401 ASARPPPGQAQ 412
```

```
US-10-106-534-2
; Sequence 2, Application US/10106534
; Patent No. US20020168668A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Millennium Pharmaceuticals, Inc.
; TITLE OF INVENTION: 14691, A Human Glutamate Receptor Family
; TITLE OF INVENTION: Member and Uses Therefor
; FILE REFERENCE: MP101-042P1RM
; CURRENT APPLICATION NUMBER: US/10/106,534
; CURRENT FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: 60/279,086
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 998
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-106-534-2

Query Match          53.3%; Score 49; DB 9; Length 998;
Best Local Similarity 66.7%; Pred. No. 69;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 STQRPQPGRQ 16
   :: |||||
Db 395 ASARPPPGQAQ 406

RESULT 5
US-09-737-149-8
; Sequence 8, Application US/09737149
; Patent No. US20020077466A1
; GENERAL INFORMATION:
; APPLICANT: Spaderna, Steven K
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Muralidhara, Padigaru
; APPLICANT: Spytek, Kimberly A.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-620 CIP
; CURRENT APPLICATION NUMBER: US/09/737,149
; CURRENT FILING DATE: 2001-06-15
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: 60/170,564
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,165
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,362
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,544
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 60/174,404
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: 60/174,962
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 60/223,929
; PRIOR FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1043
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-737-149-8

Query Match          53.3%; Score 49; DB 10; Length 1043;
Best Local Similarity 66.7%; Pred. No. 72;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 STQRPQPGRQ 16
   :: |||||
Db 401 ASARPPPGQAQ 412
```

RESULT 6
US-10-028-072-42
Sequence 42, Application US/10028072
Publication No. US20030004311A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Mei-Qiang
APPLICANT: Geriltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Matanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang
TITLE OR INVENTION:
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/028, 072
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059836
PRIOR FILING DATE: 1997-09-24
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062285
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062814
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/062816
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063045
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063082
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/063127
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063327
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063329
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063550
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063561

PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063704
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063733
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063735
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063738
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063755
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064248
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/064809
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065846
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066453
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066511
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069212
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069278
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069334
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069694
PRIOR FILING DATE: 1997-12-16
PRIOR APPLICATION NUMBER: 60/072320
PRIOR FILING DATE: 1998-01-23
PRIOR APPLICATION NUMBER: 60/073612
PRIOR FILING DATE: 1998-02-04
PRIOR APPLICATION NUMBER: 60/074086
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/074092
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-02-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081695
PRIOR FILING DATE: 1998-04-14
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081818
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082999
PRIOR FILING DATE: 1998-04-24
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07

; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085149
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086414
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086430
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088730
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088741
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090538
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07

Query Match 52.7%; Score 48.5; DB 9; Length 83;
Best Local Similarity 76.9%; Pred. No. 6.6;
Matches 10; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

OY 4 KSTQPPPPGGRQ 16
: ! ! ! ! ! ! ! !
DB 35 RSPQPPPP-GRQ 46

RESULT 7

US-10-121-049-42
; Sequence 42, Application US/10121049
; Publication No. US2003002239A1
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC17
; CURRENT APPLICATION NUMBER: US/10/121,049
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 42
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-121-049-42

Query Match 52.7%; Score 48.5; DB 9; Length 83;
Best Local Similarity 76.9%; Pred. No. 6.6;
Matches 10; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

OY 4 KSTQPPPPGGRQ 16
: ! ! ! ! ! ! ! !
DB 35 RSPQPPPP-GRQ 46

RESULT 8

US-10-123-904-42
; Sequence 42, Application US/10123904
; Publication No. US20030022328A1
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC54
; CURRENT APPLICATION NUMBER: US/10/123,904
; CURRENT FILING DATE: 2002-04-16
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 42
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-904-42

Query Match 52.7%; Score 48.5; DB 9; Length 83;
Best Local Similarity 76.9%; Pred. No. 6.6;
Matches 10; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

OY 4 KSTORPPPPGRO 16
: | | | | | | | | | |
Db 35 RSPQRP PPP-GRO 46

RESULT 9

US-10-140-470-42
; Sequence 42, Application US/10140470
; Publication No. US20030022331A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Matanabe, Collin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C160
; CURRENT APPLICATION NUMBER: US/10/140,470
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 42
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-470-42

Query Match 52.7%; Score 48.5; DB 9; Length 83;
Best Local Similarity 76.9%; Pred. No. 6.6;
Matches 10; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

OY 4 KSTORPPPPGRO 16
: | | | | | | | | | |
Db 35 RSPQRP PPP-GRO 46

RESULT 10
US-10-175-746-42
; Sequence 42, Application US/10175746
; Publication No. US20030027270A1
; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel

APPLICANT: Matanabe, Collin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C353
; CURRENT APPLICATION NUMBER: US/10/175,746
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 42
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-746-42

Query Match 52.7%; Score 48.5; DB 9; Length 83;
Best Local Similarity 76.9%; Pred. No. 6.6;
Matches 10; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

OY 4 KSTORPPPPGRO 16
: | | | | | | | | | |
Db 35 RSPQRP PPP-GRO 46

RESULT 11

US-10-176-918-42
; Sequence 42, Application US/10176918
; Publication No. US20030027275A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Matanabe, Collin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C382
; CURRENT APPLICATION NUMBER: US/10/176,918
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 42
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-918-42

Query Match 52.7%; Score 48.5; DB 9; Length 83;
Best Local Similarity 76.9%; Pred. No. 6.6;
Matches 10; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

OY 4 KSTORPPPPGRO 16
: | | | | | | | | | |
Db 35 RSPQRP PPP-GRO 46

RESULT 12
US-10-176-921-42
; Sequence 42, Application US/10176921
; Publication No. US20030027276A1

APPLICANT: Matanabe, Collin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C353
; CURRENT APPLICATION NUMBER: US/10/175,746
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 42
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-918-42


```
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C286
; CURRENT APPLICATION NUMBER: US/10/176,921
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 42
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-921-42

Query Match      52.7%; Score 48.5; DB 9; Length 83;
Best Local Similarity 76.9%; Pred. No. 6.6;
Matches 10; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 4 KSTQRPPPPQGRQ 16
Db 35 RSPQRPPPP-GRQ 46

; RESULT 13
US-10-137-865-42
; Sequence 42, Application US/10137865
; Publication No. US20030032155A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C162
; CURRENT APPLICATION NUMBER: US/10/140,474
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 42
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-474-42

Query Match      52.7%; Score 48.5; DB 9; Length 83;
Best Local Similarity 76.9%; Pred. No. 6.6;
Matches 10; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 4 KSTQRPPPPQGRQ 16
Db 35 RSPQRPPPP-GRQ 46

; RESULT 14
US-10-140-474-42
; Sequence 42, Application US/10140474
; Publication No. US20030032156A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C162
; CURRENT APPLICATION NUMBER: US/10/140,474
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 42
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-474-42

Query Match      52.7%; Score 48.5; DB 9; Length 83;
Best Local Similarity 76.9%; Pred. No. 6.6;
Matches 10; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 4 KSTQRPPPPQGRQ 16
Db 35 RSPQRPPPP-GRQ 46

; RESULT 15
US-10-142-431-42
; Sequence 42, Application US/10142431
; Publication No. US20030036179A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
```

APPLICANT: Tumas, Daniel
APPLICANT: Matanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS
FILE REFERENCE: P3330R1C251
CURRENT APPLICATION NUMBER: US/10/142,431
CURRENT FILING DATE: 2002-05-10
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 42
LENGTH: 83
TYPE: PRT
ORGANISM: Homo Sapien
US-10-142-431-42

Query Match 52.7%; Score 48.5; DB 9; Length 83;
Best Local Similarity 76.9%; Pred. No. 6.6;
Matches 10; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 4 KSTORPPPPGRO 16
:|||||
DB 35 RSPQRPPPP-GRO 46

Search completed: March 10, 2003, 18:19:49
Job time : 12.2667 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: March 10, 2003, 17:46:43 ; Search time 18 Seconds
(without alignments)
1343.372 Million cell updates/sec

Title: US-09-924-654-4

Perfect score: 3209
Sequence: 1 MSYGEISGKFLGPREETSE.....RORWGSIGSHNRVCVTNNH 583

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	154	4.8	400	1 CN37_BOVIN	P06623 bos taurus
2	148.5	4.6	421	1 CN37_HUMAN	P09543 homo sapien
3	147	4.6	420	1 CN37_MOUSE	P16330 mus musculu
4	131	4.1	1386	1 ZAP3_MOUSE	O99017 mus musculu
5	128.5	4.0	1822	1 ZAP3_HUMAN	P49750 homo sapien
6	128	4.0	964	1 YOY1_CAEEL	O09560 caenorhabdi
7	127	4.0	275	1 FRA1_RAT	P10158 rattus norv
8	124	3.9	407	1 YZRS_CAEEL	O19683 caenorhabdi
9	121.5	3.8	1059	1 CERD_RAT	P13635 rattus norv
10	119	3.7	1790	1 USOI_YEAST	P25386 saccharomyc
11	118.5	3.7	301	1 KIPN_BPT4	P06855 bacterioph
12	118	3.7	273	1 FRA1_MOUSE	P48755 mus musculu
13	117.5	3.7	952	1 IF41_YEAST	P39935 saccharomyc
14	117	3.6	1658	1 YME7_YEAST	O03661 saccharomyc
15	116	3.6	404	1 CN37_RAT	P13233 rattus norv
16	115	3.6	1164	1 KEL1_YEAST	P38853 saccharomyc
17	113	3.5	646	1 CG11_CANAL	P24866 candida alb
18	113	3.5	1157	1 SRA4_HUMAN	O95104 homo sapien
19	111	3.5	1484	1 CES2_HUMAN	O9BX13 homo sapien
20	110.5	3.4	1076	1 ARS2_DROME	O9V947 drosophila
21	110.5	3.4	1076	1 YOH8_YEAST	O08236 saccharomyc
22	110.5	3.4	1355	1 SALM_DROME	P39770 drosophila
23	110.5	3.4	1818	1 HMM2_MYCPN	P75471 mycoplasma
24	110	3.4	618	1 HMM2_MOUSE	O97052 mus musculu
25	110	3.4	645	1 MM24_HUMAN	O9Y512 homo sapien
26	110	3.4	1032	1 KINN_HUMAN	O12840 homo sapien
27	109.5	3.4	1147	1 CGAL_HELPY	P02020 helicobacte
28	109	3.4	2774	1 MAPA_RAT	P49926 rattus norv
29	108.5	3.4	271	1 FRA1_HUMAN	P15407 homo sapien
30	108	3.4	618	1 MM24_RAT	O93P46 rattus norv
31	107.5	3.3	710	1 L778_ARATH	O06738 arabidopsis
32	107.5	3.3	712	1 Y352_HUMAN	O15060 homo sapien

34	107.5	3.3	1399	1 Z291_HUMAN	O9by12 homo sapien
35	107.5	3.3	1935	1 MYSS_CYPCA	O90339 cyprinus ca
36	107.5	3.3	1939	1 MYH6_HUMAN	P13533 homo sapien
37	107.5	3.3	2869	1 RBP1_PLAVB	O00798 plasmodium
38	107	3.3	462	1 CATC_MOUSE	P97821 mus musculu
39	107	3.3	725	1 ADOB_RAT	O05764 rattus norv
40	107	3.3	731	1 BAF1_YEAST	P14164 saccharomyc
41	107	3.3	1182	1 CGA2_HELPY	P55746 helicobacte
42	107	3.3	1447	1 BUD4_YEAST	P47136 saccharomyc
43	106.5	3.3	1186	1 CAGA_HELPY	P55980 helicobacte
44	106.5	3.3	1411	1 YM42_YEAST	O03214 saccharomyc
45	106	3.3	506	1 NPL3_HUMAN	O99457 homo sapien

ALIGNMENTS

```
RESULT 1
ID      CN37_BOVIN          STANDARD;          PRT;          400 AA.
AC      P06623;
DT      01-JAN-1988 (Rel. 06, Created)
DR      01-JUL-1993 (Rel. 26, Last sequence update)
DI      30-MAY-2000 (Rel. 39, Last annotation update)
DE      2',3'-cyclic nucleotide 3'-phosphodiesterase (EC 3.1.4.37) (CNP)
DE      (CNPase).
GN      CNP.
OS      Bos taurus (Bovine).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC      Bovidae; Bovinae; Bos.
OX      NCBI_Taxid=9913;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE-Brain;
RX      MEDLINE=87137604; PubMed=3029107;
RA      Kurinara T., Fowler A.V., Takahashi Y.;
RT      "cDNA cloning and amino acid sequence of bovine brain 2',3'-cyclic-
RL      nucleotide 3'-phosphodiesterase."
RL      J. Biol. Chem. 262:3256-3261(1987).
RN      [2]
RP      REVISION TO 318.
RC      TISSUE-Brain;
RA      Kurinara T., Fowler A.V., Takahashi Y.;
RT      J. Biol. Chem. 262:16754-16754(1987).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      TISSUE-Retina;
RX      MEDLINE=88015560; PubMed=2821502;
RA      Vogel U.S., Thompson R.J.;
RT      "Nucleotide sequence of bovine retina 2',3'-cyclic nucleotide 3'-
RL      phosphodiesterase."
RL      Nucleic Acids Res. 15:7204-7204(1987).
RN      [4]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=87247281; PubMed=3036592;
RA      Vogel U.S., Thompson R.J.;
RT      "Molecular cloning of the myelin specific enzyme 2',3'-cyclic-
RL      nucleotide 3'-phosphodiesterase."
RL      FEBS Lett. 218:261-265(1987).
RN      -1- CATALYTIC ACTIVITY: Nucleoside 2',3'-cyclic phosphate + H(2)O -
CC      nucleoside 2'-phosphate.
RN      -1- SUBCELLULAR LOCATION: FIRMLY BOUND TO MEMBRANE STRUCTURES OF BRAIN
CC      WHITE MATTER.
RN      -1- PTM: MET-1 MAY BE REMOVED AFTER TRANSLATION.
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
```

```
CC -----
DR EMBL; J02659; AAA30456.1; ALT_SEQ.
DR EMBL; Y00405; CAAG6466.1;
DR EMBL; M27606; AAA30457.1;
DR PIR; A26861; ESBOP3.
DR InterPro; IPR001230; Prenyl_site.
KW Hydrolase; Membrane; Brain.
SQ SEQUENCE 400 AA; 44875 MW; 3448FC367D647CF8 CRC64;

Query Match 4.8%; Score 154; DB 1; Length 400;
Best Local Similarity 28.6%; Pred. No. 0.0092;
Matches 44; Conservative 37; Mismatches 49; Indels 24; Gaps 5;

QY 394 EEKLNKIQ--KLILLRGLPGSGKTTLSRILLGQNRDGVFTSDDDYFHHQDGYRYNVNQL 451
DB 20 ETVATLQCKFLFIRLGLPGSGKSTLAFIVDKYRDGKMYVSADSKYKTTPGARGSFS- 78

QY 452 GDAHDNQNRKAOIDGRSP-----VIIDNTNIOAWEMKPYVEVAIGKGYRVEFHE 503
DB 79 ----EYKQ-----LDELAACCRDRFVLVLDNTNHERERLEQLFELADQYQYQVVLVE 128

QY 504 PETWKKFDEEELKRNKHGVSR---KKIAQMLDR 534
DB 129 PKTAWRLDCAQLKERNQWOLSADLKKLPGLEK 162

RESULT 2
CN37_HUMAN STANDARD; PRT; 421 AA.
AC P09543;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 2',3'-cyclic nucleotide 3'-phosphodiesterase (EC 3.1.4.37) (CNP)
DE (CNase).
GN CNP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93050745; PubMed=1385234;
RA Thompson R.J.;
RT "2',3'-cyclic nucleotide-3'-phosphohydrolase and signal transduction
RT in central nervous system myelin.";
RL Biochem. Soc. Trans. 20:621-626(1992).
RN [2]
RP SEQUENCE FROM N.A. (SHORT FORM).
RC TISSUE=Brain;
RX MEDLINE=88209067; PubMed=2835044;
RA Kurihara T., Takahashi Y., Nishiyama A., Kumanishi T.;
RT "cDNA cloning and amino acid sequence of human brain 2',3'-cyclic-
RT nucleotide 3'-phosphodiesterase.";
RL Biochem. Biophys. Res. Commun. 152:837-842(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=93314977; PubMed=8392017;
RA Monoh K., Kurihara T., Takahashi Y., Ichikawa T., Kumanishi T.,
RA Hayashi S., Minoshima S., Shimizu N.;
RT "Structure, expression and chromosomal localization of the gene
RT encoding human 2',3'-cyclic nucleotide 3'-phosphodiesterase.";
RL Gene 129:297-301(1993).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=93080285; PubMed=1360194;
RA Douglas A.J., Fox M.F., Abbott C.M., Hinks L.J., Sharpe G.,
RA Povey S., Thompson R.J.;
RT "Structure and chromosomal localization of the human 2',3'-cyclic
RT nucleotide 3'-phosphodiesterase gene.";
RL Ann. Hum. Genet. 56:243-254(1992).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
```

```
RC TISSUE=Brain, and Skin;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Nucleoside 2',3'-cyclic phosphate + H(2)O -
CC nucleoside 2'-phosphate.
CC -!- SUBCELLULAR LOCATION: FIRMLY BOUND TO MEMBRANE STRUCTURES OF BRAIN
CC WHITE MATTER.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS: CNP1/DNAI AND CNP1/DNAI
CC (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S46849; AAB23928.2;
DR EMBL; S46843; AAB23928.2; JOINED.
DR EMBL; S46845; AAB23928.2; JOINED.
DR EMBL; S46846; AAB23928.2; JOINED.
DR EMBL; M19650; AAA35704.1;
DR EMBL; D13146; BAA39694.1;
DR EMBL; D13144; BAA39694.1; JOINED.
DR EMBL; D13145; BAA39694.1; JOINED.
DR EMBL; D13146; BAA02435.1;
DR EMBL; D13144; BAA02435.1; JOINED.
DR EMBL; D13145; BAA02435.1; JOINED.
DR EMBL; S50017; AAB24298.2;
DR EMBL; S50013; AAB24298.2; JOINED.
DR EMBL; S50014; AAB24298.2; JOINED.
DR EMBL; S50016; AAB24298.2; JOINED.
DR EMBL; BC001362; AAH01362.1;
DR EMBL; BC006392; AAH06392.1;
DR EMBL; BC011046; AAH11046.1;
DR EMBL; BC028040; AAH28040.1;
DR PIR; A27703; A27703.
DR PIR; JCI518; JCI518.
DR Gene; HGNC:2158; CNP.
DR MIN; 123830;
DR InterPro; IPR001230; Prenyl_site.
KW Hydrolase; Membrane; Brain; Alternative splicing.
FT VARSPPLIC 1 20 MISSING (IN ISOFORM CNP1).
SQ SEQUENCE 421 AA; 47578 MW; CA6D0097DFD87255 CRC64;

Query Match 4.6%; Score 148.5; DB 1; Length 421;
Best Local Similarity 28.7%; Pred. No. 0.022;
Matches 39; Conservative 36; Mismatches 56; Indels 5; Gaps 2;

QY 402 KLLILLRGLPGSGKTTLSRILLGQNRDGVFTSDDDYFHHQDGYRYNVNQLGDAHDNQNR 461
DB 50 KTLFIRLGLPGSGKSTLAFIVDKYRDGKMYVSADSKYKTTPGARGSFS--EYKRLDEDL 107

QY 462 AKQAIQGRSPVIIDNTNIOAWEMKPYVEVAIGKGYRVEFHEPETWKKFDEEELKRNKH 521
DB 108 AAYCRRDRIRILVLDNTNHERERLEQLFELADQYQYQVVLVEPKTAWRLDCAQLKKNOW 167

QY 522 GVSR---KKIAQMLDR 534;
DB 168 QLSADDLKKLPGLEK 183

RESULT 3
CN37_MOUSE STANDARD; PRT; 420 AA.
AC P16330;
DT 01-AUG-1990 (Rel. 15, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 2',3'-cyclic nucleotide 3'-phosphodiesterase (EC 3.1.4.37) (CNP)
DE (CNase).
GN CNP OR CNP1.
```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM CNPI).
RX MEDLINE=90121227; Pubmed=2558653;
RA Monoh K., Kurihara T., Sakimura K., Takahashi Y.;
RT "Structure of mouse 2',3'-cyclic-nucleotide 3'-phosphodiesterase
RT gene";
RL Biochem. Biophys. Res. Commun. 165:1213-1220(1989).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS CNPI AND CNPII).
RX MEDLINE=90358801; Pubmed=2167669;
RA Kurihara T., Monoh K., Sakimura K., Takahashi Y.;
RT "Alternative splicing of mouse brain 2',3'-cyclic-nucleotide 3'-
RT phosphodiesterase mRNA";
RL Biochem. Biophys. Res. Commun. 170:1074-1081(1990).
CC -1- CATALYTIC ACTIVITY: Nucleoside 2',3'-cyclic phosphate + H2O =
CC nucleoside 2'-phosphate.
CC -1- SUBCELLULAR LOCATION: FIRMLY BOUND TO MEMBRANE STRUCTURES OF BRAIN
CC WHITE MATTER.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: CNPI/DNAI AND CNPII/DNAII
CC (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M31810; AAA37429.1; -
DR EMBL: M58045; AAA37430.1; -
DR EMBL: D38642; BAA07621.1; JOINED.
DR EMBL: D38640; BAA07621.1; JOINED.
DR EMBL: D38641; BAA07621.1; JOINED.
DR EMBL: D38642; BAA07622.1; -
DR EMBL: D38640; BAA07622.1; JOINED.
DR EMBL: D38641; BAA07622.1; JOINED.
DR PIR: A35708; ESM332.
DR MGI: MGI:88437; Cnpi.
DR InterPro: IPR001230; Prenyl-site.
KW Hydrolyase; Membrane; Brain; Alternative splicing.
FT VAAPSLIC 1 20 MISSING (IN ISOFORM CNPI).
FT CONFLICT 115 115 M -> I (IN REF. 1).
FT CONFLICT 136 136 L -> Y (IN REF. 1).
SQ SEQUENCE 420 AA; 47123 MW; DD96FED47AD15D3 CRC64;
Query Match 4.6%; Score 147; DB 1; Length 420;
Best Local Similarity 27.8%; Pred. No. 0.027;
Matches 40; Conservative 34; Mismatches 48; Indels 22; Gaps 3;
QY 402 KLLILRLPSGKTTLSRIILGONRGIVSTDDYFHHDGGRYNNQLDADHANNOR 461
DB 50 KTLRLRLGPSGKSTLRLILEKRYHDGTMKVSADAYKIIPSGRDFSE----- 98
QY 462 AKQALDGRS-----PVIIDNTNQAMKRYVEVALIKGYRVEHEDETMKFPDE 513
DB 99 AYKRLDELDLAYCRDMKRVLLDDTNNHERRLDOLFEILDQYQYOVVLEFKTAVRLDCA 158
QY 514 ELEKRNKHGVR--KRIQMIDR 534
DB 159 QLKERNQWLSADDLKRLKPGLEK 182
RESULT 4
ZAP3_MOUSE STANDARD; PRT; 1386 AA.
AC O9R017;
DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nuclear protein ZAP3.
GN ZAP3 OR ZAP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal liver;
RA Misawa K., Nosaka T., Kitamura T.;
RT "A huge nuclear protein rich in proline similar to human hypothetical
RT protein zap3 and zap13";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB033168; BAA85182.1; -
DR MGI: MGI:1926195; Zap3.
KW Nuclear protein.
FT DOMAIN 15 204 PRO-RICH.
FT DOMAIN 355 473 GLN-RICH.
FT DOMAIN 925 1012 ARG-RICH.
SQ SEQUENCE 1386 AA; 155130 MW; D862F9918ED221DF CRC64;
Query Match 4.1%; Score 131; DB 1; Length 1386;
Best Local Similarity 17.1%; Pred. No. 1.2;
Matches 121; Conservative 78; Mismatches 179; Indels 328; Gaps 28;
QY 91 PDVIESIDSOVLQARPPVLSADEIYTSKAFIPIYKPEKKRN-----EG 139
DB 612 PPAAGSONSQIPEKPRALLPTPVFSFGSTPS--PYHPPOSQVNSKPLNKVFSSBOG 668
QY 140 RNEAHVLN-----GI-----NRRGGQKQKQKQKSEKSEIDNELFQ 174
DB 669 LGESSALSQSTIAAKDTPVKSGLLADPPKGSFLGPRGPREQR----- 712
QY 175 FYKEIELEKEXDEGENSKSEPSOQFVPEYGHNNGLKLPDEKKDLSNKM-PSHC 233
DB 713 -----EQLQKAKD-----FGSEFQMAHLR-----PPDSRLQNPSPRGVPPPG 751
QY 234 DYQ--QNLGNPEPDKYPCNGQVITPFCDTSTFSF---RPMQSVYPIYVYPPPLPSLNYH 288
DB 752 SYRPPPPGKPP-----GSIVRPSAPPARRSSIPMTRP-----PVPIPPPPPPP 795
QY 289 LNIIRFSGPPRPNINPQADDSQIQNGXYVNNCHVMNMTFPPONNETYDCSENR--- 344
DB 796 -----PPPPPPVYKSTSSVKQERKWDSEFGLM-----DTNDQGLNSEKRPDTA 842
QY 345 -----SSVHPSG----- 351
DB 843 TTPSAPVLPPEPVVSSITPPPGPMGMPKPPPVQHTVDYGHGRDPTKVKVEQIPYE 902
QY 352 -----NGCSMDQRY-----VSN-----GCEVRE----- 370
DB 903 RITLRPPLPERSTFDADAGORRYRDRDRREYFPDRPSNITDRDFKRDRETHRRDR 962
QY 371 -----RCKWDHCKMKNHGTDR-----FV 388
DB 963 DVALDYERDRDRERRPRDDNOSTROKDDHSSSRGSGFDRPSYDRKSDRPYGPFG 1022
QY 389 NQO--FOEKL-----NKLOKLLILRLGL 410
DB 1023 GERTIYPERMRPLAPALGHQPPPVPRVEKKPEKKNVDILKPPGRSRRPRIYVIMKGL 1082

or send an email to license@sb-stb.ch.

CC EMBL: Z47069; CAAB738.1; .
 DR WormPep; F36G3.1; CE15979.
 KW Hypothetical protein; Coiled coil.
 FT DOMAIN 656 840 COILED COIL (POTENTIAL).
 SQ SEQUENCE 964 AA; 10866 MW; DABA6BE59350F076 CRC64;

Query Match 4.0%; Score 128; DB 1; Length 964;
 Best Local Similarity 20.5%; Pred. No. 1.2;
 Matches 96; Conservative 76; Mismatches 171; Indels 126; Gaps 23;

QY 2 SYGIE-CKRLGPREETSPRCKKLKSTESYVFNHNSNADFRIDEKGNMVPYTI 60
 DB 227 SFGSSDGDHLEVKSEATNPE--YNTSEF-----EVKNDIVYEVN 268
 QY 61 DVRGHSYLEENKIKTT--DLHRPLHDEMPGNRPVIESIDQVLOEARPLVSADEIYS 118
 DB 269 N-----ESTIDSVVIDSGTPTKDKR--NEVNISQNVNVLVRLESH--LENYDDE--T 314
 QY 119 TSKAFIPRIKPKPKKRNENEAHVLNGINDRGQKE--KQFNSKSEIDNELFOFY 176
 DB 315 TMTVIAECVYDAQENKCE--KNEEDLSPVLDENKIKNSPEKFCFRAE---SECY 367
 QY 177 KEIIELEKENDGFENSKESPEQOFVPEEGHNNGLKPKDEKKNLSKAMPSHDDYQ 236
 DB 368 EQTEETDOIK-----QFVPLME-----VSANQDFELCDIE 399
 QY 237 ONLNEPDKYVCNGQVIFPCDTSFTSFRPEM--QSVYPTVPGPLPSNHYHLNIQRF 294
 DB 400 RNGNSNDK-----PKDLNENKAFDAEHENSYLEY-----EQLPVFSSETTAKT 446
 QY 295 SGPNPSPNIFQADDSQ-----IQNGY--YVNNCHVNMCMTEPDQNNETDCSEN 344
 DB 447 PIPVHEHSISGVKESQSVFNOAQNELMPTHTIISDEVTSTADNNIPFSDCSTHK 506
 QY 345 SSV---HPSNGSGMODRYNSNCFCEVRECKMDHCMDK-----NGTD- 385
 DB 507 VKVDFDGISSGKEBSLSSSTSESF-EFAEPVTENHIODERAQALRNSIIRYPSNDTD 565
 QY 386 ---HFVNOQFOEKL-----NKLQKLLILRLPGSGKTLRSHLLQ 425
 DB 566 DELDSVGEFDEDLIAVKQISAEVQLVAALNAGRDEQMSAVMWGK 614

RESULT 7
 FRAL RAT STANDARD; PRT; 275 AA.

ID P10158; .
 AC 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fos-related antigen 1 (FRA-1).
 GN FOSL1 OR FRA1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN 11
 RP SEQUENCE FROM N.A.
 RA MEDLINE=88261282; PubMed=1133553;
 RX Cohen D.R., Curran T.,
 RT "fra-1: a serum-inducible, cellular immediate-early gene that encodes
 a fos-related antigen";
 RL Mol. Cell. Biol. 8:2063-2069(1988).
 RN 121
 RP SEQUENCE OF 1-43 FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RA MEDLINE=95311973; PubMed=7791782;
 RX Bergers G., Gränlinger P., Braselmann S., Wrighton C., Busslinger M.,
 RT "transcriptional activation of the fra-1 gene by AP-1 is mediated by
 regulatory sequences in the first intron";
 RL Mol. Cell. Biol. 15:3748-3756(1995).

CC -1 SUBUNIT: HETERODIMER (BY SIMILARITY).
 CC -1 SUBCELLULAR LOCATION: Nuclear.
 CC -1 INDUCTION: SERUM-INDUCIBLE.
 CC -1 SIMILARITY: BELONGS TO THE BZIP FAMILY. FOS SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-stb.ch/announce/>
 CC or send an email to license@sb-stb.ch).

CC EMBL: M19651; AAA4117.1; .
 DR EMBL: U24154; AAA82045.1; .
 DR PIR: A27722; TVRTER.
 DR HSSP: P01100; 1FOS.
 DR TRANSFAC: T01208; .
 DR InterPro: IPR000837; Leuzip_Fos.
 DR InterPro: IPR004827; TF_bzip.
 DR Pfam: PF001070; bzip.1; Leuzip_Fos.
 DR PRINTS: PR00042; LEUZIPRPOS.
 DR SMART: SM00338; BRLZ.1.
 DR PROSITE: PS00036; BZIP_BASIC; 1.
 KW Nuclear protein; DNA-binding.
 FT DOMAIN 105 108 POLY-GLU.
 FT DVA_BIND 113 131 BASIC MOTIF.
 FT DOMAIN 135 163 LEUCINE-ZIPPER.
 FT FT DOMAIN 255 259 POLY-SER.
 SQ SEQUENCE 275 AA; 30115 MW; 103726AD5D1FAB2F CRC64;

Query Match 4.0%; Score 127; DB 1; Length 275;
 Best Local Similarity 22.9%; Pred. No. 0.28; Indels 88; Gaps 13;
 Matches 62; Conservative 30; Mismatches 91;

QY 90 RPDVIESIDQVLOEAR---PPL--VSADDEI-YTSKAFIGPI-----YRP- 130
 DB 20 RPAQPOAQOTVQOQKFLHVPISINAVSGSELQMMQPHLGSGVPRPLTYQSPQ 79
 QY 131 -----PEKKRNEGNEAHVLNGINDRGQKEKQNF-NSEK 165
 DB 80 PRPGVIRALGPPRRRRPQOISPEEERRRVRERRNKLAAKCRNRKELTDFLOAET 139
 QY 166 SEINNELFQFYKTEIELEKEDGFENSKESPEQOQFVPEEGHNNGLKPKDEKKDLS 225
 DB 140 DKLEDEKSGLOREIELOKQERLE-----LVLEAHRPICKIPEEDKDTG 185
 QY 226 NKAMPSHCDYQONLGNEDPKYPCN-----GQVI-----PFCDDT-SFTSFRPEW 268
 DB 186 GTSSTS-----GAGSPRG--PCRPVPCISLSPGVLEPFLALHTLTPTSLPFPSTL 237
 QY 269 QSVYPTVPGPLPSLNYHLNIQRFSGPPN 299
 DB 238 VFTYF-----STPEPCSSAHRKSSSSGSDPS 263

RESULT 8
 YZRS CAEEL STANDARD; PRT; 407 AA.

ID YZRS CAEEL .
 AC 019683; O9G064;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein F21D5.5 in chromosome IV.
 GN F21D5.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 CC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Berks M.;

```

RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Durbin R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Kohara Y., Shin'ichi T., Suzuki Y., Sugano S., Potdevin M.,
  Thierry-Mieg Y., Thierry-Mieg D., Thierry-Mieg J.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO YEAST YMR156C, S.POMBE SPAC23C11.04C AND
  ACNVPV ORF33.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/ or
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 254271; CAA91035.2; -
DR EMBL; AF292046; AAG41142.1; -
DR WormPep; F21D5.5; CE27135.
KW Hypothetical protein.
SQ SEQUENCE 407 AA; 46188 MW; A17DCAFCB979EF3C CRC64;
-----
Query Match 3.9%; Score 124; DB 1; Length 407;
Best Local Similarity 19.8%; Pred. No. 0.71;
Matches 78; Conservative 55; Mismatches 130; Indels 130; Gaps 16;
QY 182 LSEKDGFGNSCKESEPSEQEVPF-----YEGHNNGLLKDEKDLKSNKAMPCHCDVQ 236
DB 1 MKRAMDGAKNEAKSKDPTLFGRLPKKTGSWESKNGDLMIFTH-----SECEGK 50
QY 237 QNLGNEDPKYPCNGOVIPTFCDTSTSPFQWQSVYPTVPYGPPLPSLNYHL----- 289
DB 51 EKIA----AFDMGLIKTKSKVPTNCQDWQLYDSIPSPFKLHSDGFKIVFTNQK 106
QY 290 -----NIQFSGPNPSPNIFQADDSQIQNGYVNNCHVNNCMTFDQNN 335
DB 107 GIHAGKVDNREFKRIEATVGLGIPVQAFVS-----VAGHYRKPVCVGMNELKL--RN 159
QY 336 EYTDSCENRSSHVPSNGSCMDRIVSN--GFCVERCKWCHDKMDKNGTDRF-----VN 389
DB 160 DEVEINEKESI-----FVGDAAGRIKTTSPKKD-----HSYADRFFAANGV 201
QY 390 QOQF-----EELKNKLOKLLILLRLPGS 413
DB 202 VKFQTEPEFFGSKVDPMGPNFDPKNLFSEITELEPHDAQLSKSEKEILLMVGFPQS 261
QY 414 GTTTSRILLGNRDGIVFSTDYFHQDGYR-YVNNQLGDADHNNQNR--KQAIQDGR 470
DB 262 GKSTFAKML-----GHQHDYKIVNRDTIG---TWQCKVAATRSYLDCK 302
QY 471 SPVIDNTNIQAWENKPYVEVAIGKGYRVEFHE 503
DB 303 S-WIDNTSPDLRSKRKYIDVAKELGVPIRCFE 334
-----
RESULT 9
CERU_RAT
ID CERU_RAT STANDARD: PRT; 1059 AA.
AC P13635; Q64719;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ceruloplasmin precursor (EC 1.16.3.1) (Ferroxidase).
GN CP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE=Lung, and Liver;
RA MEDLINE=90237081; PubMed=2332446;
RX Fleming R.E., Gitlin J.D.;
RT "Primary structure of rat ceruloplasmin and analysis of
  tissue-specific gene expression during development.";
RL J. Biol. Chem. 265:7701-7707(1990).
RN [2]
RP SEQUENCE OF 257-294; 571-612 AND 823-892 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=87137545; PubMed=3818625;
RA Aldred A.R., Grimes A., Schreiber G., Mercer J.F.B.;
RT "Rat ceruloplasmin. Molecular cloning and gene expression in liver,
  choroid plexus, yolk sac, placenta, and testis.";
RL J. Biol. Chem. 262:2875-2878(1987).
CC -!- FUNCTION: CERULOPLASMIN IS A BLUE, COPPER-BINDING (6-7 ATOMS PER
  MOLECULE) GLYCOPROTEIN FOUND IN PLASMA. FOUR POSSIBLE FUNCTIONS
  ARE FERROXIDASE ACTIVITY, AMINE OXIDASE ACTIVITY, COPPER TRANSPORT
  AND HOMEOSTASIS, AND SUPEROXIDE DISMUTASE ACTIVITY.
CC -!- FUNCTION: MAY ALSO PLAY A ROLE IN FETAL LUNG DEVELOPMENT OR
  PULMONARY ANTIOXIDANT DEFENSE.
CC -!- CATALYTIC ACTIVITY: 4 Fe(2+) + 4 H(+) + O(2) = 4 Fe(3+) + 2 H(2)O.
CC -!- COFACTOR: BINDS 6 CU-IONS PER MOLECULE. THIS PROTEIN BELONGS TO
  THE MULTICOPPER OXIDASES WHICH CONTAIN THREE DISTINCT CU CENTERS
  KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE 3 OR COUPLED
  BINUCLEAR.
CC -!- TISSUE SPECIFICITY: SYNTHESIZED IN LIVER AND SECRETED INTO THE
  PLASMA. ALSO CHOROID PLEXUS, YOLK SAC, PLACENTA, AND TESTIS; NOT
  IN STOMACH AND SMALL INTESTINE. FETAL LUNG AND LIVER.
CC -!- INDUCTION: BY INFLAMMATION.
CC -!- SIMILARITY: CONTAINS 3 F5/8 TYPE A DOMAINS; EACH IS COMPOSED OF
  2 PLASTOCYANIN-LIKE REPEATS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/ or
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L33869; AAA40917.1; -
DR EMBL; M80529; AAB65820.1; -
DR EMBL; J02670; AAA40914.1; ALT_SEQ.
DR EMBL; M14102; AAA40915.1; -
DR PIR; A35210; A35210
DR PIR; A29564; A29564.
DR HSP; P00450; ILCW.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR002355; MultiCu-oxidase2.
DR Pfam; PF00394; Cu-oxidase; 3.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
DR PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
DR Oxidoreductase; Copper; Metal-binding; Glycoprotein; Plasma; Repeat;
KW Signal.
FT SIGNAL 1 19 PROBABLE.
FT CHAIN 20 1059 CERULOPLASMIN.
FT DOMAIN 20 356 F5/8 TYPE A 1.
FT DOMAIN 20 199 PLASTOCYANIN-LIKE 1.
FT DOMAIN 208 354 PLASTOCYANIN-LIKE 2.
FT DOMAIN 369 712 F5/8 TYPE A 2.
FT DOMAIN 564 710 PLASTOCYANIN-LIKE 3.
FT DOMAIN 724 1055 PLASTOCYANIN-LIKE 4.
FT DOMAIN 724 894 F5/8 TYPE A 3.
FT DOMAIN 902 1051 PLASTOCYANIN-LIKE 5.
FT DISULFID 173 199 BY SIMILARITY.
FT DISULFID 275 356 BY SIMILARITY.
FT DISULFID 528 554 BY SIMILARITY.
FT DISULFID 631 712 BY SIMILARITY.
FT DISULFID 868 894 BY SIMILARITY.
```



```
QY 15 EEVTS-EPRCKLK-----STTESYVFHNSNA-----DFHRIQEKTKGNDWVPV 57
DB 730 EEVEKLQROCTKLKGEITSLOQTETES-----THENLTEKLIALTNEHKELDEK-----Y 778

QY 58 TIIDVRGHYSLOEN-KIKTTDLH--RPLHDEMPGNRPDVIESIDQ---VLOEARPLVS 111
DB 779 QILN--SSHSLSKENSILTEKUNVRDSLDMTQUR-DVLETKDKENOTALLEYKSTIHK 836

QY 112 ABDEIYSTKAFIGPIYKPKKKRNEGRNE-----AHVLNGT--NDRGGQKEKQKF 161
DB 837 QEDSIKTEKLETL--SOKKKAEDGINKMGKDLFALSREMQAVEENCKNLQKEKOKS 893

QY 162 N-----SEKSEIDNELFQYKEIELEKEKDFENSCKSESPSQOFPVF-----YEGH 210
DB 894 NYNHOKETKSLKREDIAAKTEIKAINENLEEMKIQCNLSKEKEHISKELVELYKSRFOSH 953

QY 211 NNGLLKPDEKDLN 226
DB 954 DNLVAKLTEKLSLAN 969

RESULT 11
ID KIPN_BPT4 STANDARD; PRT; 301 AA.
AC P06855;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Polynucleotide kinase (EC 2.7.1.78) (PNK) (Polynucleotide 5'-hydroxyl-
DE kinase).
GN PSET.
OS Bacteriophage T4.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC T4-like viruses.
OX NCBI_TaxID=10665;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86030251; PubMed=2996886;
RA Midgley C.A., Murray N.E.;
RT "T4 polynucleotide kinase; cloning of the gene (pset) and
RT amplification of its product.";
RL EMOB J. 4:2695-2703(1985).
RN [2]
RP SEQUENCE FROM N.A.
RA Kutter E., Arisaka F., Kunisawa T., Tsugita A., Mosig G.,
RA Mesyanzhinov V., Ruger W., Stidham T., Thomas E.;
RT "Bacteriophage T4 genome analysis.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ACTS BOTH AS A 5'-KINASE AND A 3'-PHOSPHATASE. CATALYZES
CC THE TRANSFER OF THE TERMINAL PHOSPHATE OF ATP TO THE 5'-HYDROXYL
CC TERMINI OF RIBO- AND DEOXYRIBONUCLEOTIDES. IN THE PRESENCE OF ADP
CC THE ENZYME ALSO CATALYZES AN EXCHANGE REACTION. IN THE EXCHANGE
CC REACTION, AN EXCESS ADP CAUSES THE ENZYME TO TRANSFER THE 5'
CC TERMINAL PHOSPHATE FROM PHOSPHORYLATED DNA TO ADP.
CC -!- CATALYTIC ACTIVITY: ATP + 5'-dephospho-DNA = ADP + 5'-phospho-DNA.
CC -!- DATABASE: NAME=Worthington enzyme manual;
CC WWW="http://www.worthington-biochem.com/manual/P/PNKT4.html".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X03007; CRA26792.1; -.
CC DR EMBL; AF158101; AAD42642.1; -.
CC DR PIR; A24642; KIBPP4.
CC KW Transferase; Kinase.
CC SQ SEQUENCE 301 AA; 34620 MW; 0EAFB6BA83236D31 CRC64;
```

```
Query Match 3.7%; Score 118.5; DB 1; Length 301;
Best Local Similarity 30.5%; Pred. No. 1.1;
Matches 46; Conservative 18; Mismatches 60; Indels 27; Gaps 7;

QY 402 KLLILLRLGPGSGKTTLSRLLGQNRDGIIVFSTDDY-----FHHQDGYRYNNVNGDAH 455
DB 2 KRIIITICPGSGKSTWAREFTAKNPGFYNNRDYROSIMAAHEERDEYKTKKREGIVT 61

QY 456 DWNQNRKAQADQGRS--PVIIDNTNIQ-----AME--MKPYVEVAIGKGVREVEHEPET 506
DB 62 GMPDTAKSILYGGDSVKGVIISDNLNPERRLAWETFAKEY-----GKVKVEHKVEDV 114

QY 507 WWKFPDEBELEKRN-KHGVSRKKIAQMDLRYE 536
DB 115 PW----TELVRNRSKRGTKAVPIDVLRSMYK 141

RESULT 12
FRAL_MOUSE STANDARD; PRT; 273 AA.
ID FRAL_MOUSE
AC P48755; O35285;
DT 01-FEB-1996 (Rel. 33, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fos-related antigen-1 (FRA-1).
GN FOSL1 OR FRA1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/CBYJ; TISSUE=Spleen;
RX MEDLINE=97047769; PubMed=8892610;
RA Huo L., Rothstein T.L.;
RT "Isolation and characterization of murine fra-1: induction mediated
RT by CD40 and surface Ig is protein kinase C dependent.";
RL J. Immunol. 157:3812-3818(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RX MEDLINE=97440303; PubMed=9294610;
RA Schreiber M., Poirier C., Franchi A., Kurzbauer R., Guenet J.L.,
RA Carle G.F., Wagner E.F.;
RT "Structure and chromosomal assignment of the mouse fra-1 gene, and its
RT exclusion as a candidate gene for oc (osteosclerosis).";
RL Oncogene 15:1171-1178(1997).
CC -!- SUBUNIT: HETERODIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE BZIP FAMILY. FOS SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U34245; AAC52888.1; -.
CC DR EMBL; AF017128; AAB71369.1; -.
CC DR HSSP; P01100; 1FOS.
CC DR TRANSFAC; T00292; -.
CC DR MGD; MGI:107179; Fosl1.
CC DR InterPro; IPR000837; Leuzip_Fos.
CC DR InterPro; IPR004827; TF_BZIP.
CC DR Pfam; PF00170; bzip; 1.
CC DR PRINTS; PR00042; LEUZIPPREFOS.
CC DR SMART; SM00338; BRLZ; 1.
CC DR PROSITE; PS00036; BZIP_BASIC; 1.
CC KW Nuclear protein; DNA-binding.
CC FT DOMAIN 103 106 POLY-GLU.
CC DNA_BIND 111 129 BASIC MOTIF.
CC FT
```

```

FT DOMAIN 133 161 LEUCINE-ZIPPER.
FT DOMAIN 253 257 POLY-SER.
FT CONFLICT 17 17 P -> A (IN REF. 1).
FT CONFLICT 39 39 L -> F (IN REF. 1).
FT CONFLICT 170 170 P -> L (IN REF. 1).
FT CONFLICT 249 249 A -> T (IN REF. 1).
SQ SEQUENCE 273 AA; 29799 MW; 48068B0259866805 CRC64;

Query Match 3.7%; Score 118; DB 1; Length 273;
Best Local Similarity 22.5%; Pred. No. 1;
Matches 60; Conservative 27; Mismatches 104; Indels 76; Gaps 11;

QY 87 PGNRPDVEISDSQVLOAR-----PVLASADEI-YNSKAFIGPI-----YKP-130
DB 17 PYGRAPPOAOAQTAAQOKFHLVPSIDSSOELHWMYQPHLGTGTPRPLAYQYSP 76
QY 131 -----PEKKRNEGRNEAHVINGIDRGOKOKF-NSE 164
DB 77 QPRGVIALGPPGVRRRPCEQISPEDEERRVRRENKLAACGRNRKELTDFLQAE 136
QY 165 KSEIDNELFQYKTEIELEKEKGFENSCKESEPOFVPEYEGHNNGLKPDDEKDL 224
DB 137 TDKLEDEKSGLOREIEELQOKERLE-----LYLEAHRPICKIPEDDKDP 182
QY 225 SNKAMPSHCDYQONNGNEPDYPC-----NGQYI-----PFCDF-SPTSPRPMQSY 272
DB 183 GSGGSTSASSPPAPGR---PVPCISLSPGVLPEALHTPLMTPTSLTPPSLVET 239
QY 273 PFIYVPGPLPSLNYHLNIORFSGPPN 299
DB 240 P-----SPEPSSAHARRKSSSSGDP 261

RESULT 13
IF41_YEAST
ID IF41_YEAST STANDARD; PRT; 952 AA.
AC P39935.
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Eukaryotic initiation factor 4F subunit p150 (eIF-4F) (mRNA cap-binding protein complex subunit p150).
GN TIF4631 OR YGRL62W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93330281; PubMed=8336723;
RA Goyer C., Altman M., Lee H.S., Blanc A., Deshmukh M., Woolford J.L., Trachsel H., Sonenberg N.;
RT "TIF4631 and TIF4632: two yeast genes encoding the high-molecular-weight subunits of the cap-binding protein complex (eukaryotic initiation factor 4F) contain an RNA recognition motif-like sequence and carry out an essential function."
RL Mol. Cell. Biol. 13:4860-4874(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C;
RA MEDLINE=97435481; PubMed=9290212;
RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae chromosome VII."
RT Yeast 13:1077-1090(1997).
CC -I- FUNCTION: INTERACTS WITH THE MRNA CAP STRUCTURE, M7GPPX. THIS INTERACTION IS REQUIRED FOR EFFICIENT RIBOSOME BINDING TO THE MRNA. TIF4631 IS PROBABLY ESSENTIAL WHEN TIF4632 IS MISSING.
CC -I- SUBUNIT: THE CAP-BINDING PROTEIN COMPLEX IS COMPOSED OF AT LEAST TWO PROTEINS, A 24 KDA (TIF45) AND A 150-200 KDA SUBUNIT (TIF4631).
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L16923; AAA02757.1;
DR EMBL: Z72947; CA97184.1;
DR PIR: A48086; A48086.
DR SGD: S0003394; TIF4631.
DR InterPro: IPR003890; IF_eIF4G.
DR Pfam: PF02854; MIF4G; 1.
DR SMART: SM00543; MIF4G; 1.
KW Initiation factor; Protein biosynthesis; RNA-binding;
KW Multigene family.
FT DOMAIN 173 200 PRO/SER-THR-RICH.
FT DOMAIN 375 384 ALA/GLU-RICH.
FT DOMAIN 488 553 ARG/SER-RICH.
FT DOMAIN 873 899 ARG/SER-RICH.
FT DOMAIN 869 872 POLY-GLU.
FT CONFLICT 7 7 H -> Q (IN REF. 1).
FT CONFLICT 37 37 T -> N (IN REF. 1).
FT CONFLICT 110 110 Q -> K (IN REF. 1).
FT CONFLICT 207 207 R -> K (IN REF. 1).
FT CONFLICT 361 361 D -> E (IN REF. 1).
SQ SEQUENCE 952 AA; 107101 MW; 391256802F86118E CRC64;

Query Match 3.7%; Score 117.5; DB 1; Length 952;
Best Local Similarity 19.5%; Pred. No. 5.3;
Matches 112; Conservative 85; Mismatches 217; Indels 159; Gaps 29;

QY 81 PLHDEPGNRPDVEISDSQ-----VLOARPPVLSADE-IYSTKATIG 125
DB 423 PIEDVFSFNYPEGIEGPDIKYKHEVKYTGTFLOEKDLNVKADAEWVOSTASKIVI 482
QY 126 PIYKPEKKKRNREG-----NEAH-----VLNGIDRGOKOKGFENSEKSIDNELOF 175
DB 483 PEGMGRNRSRSGFGNNSSGHDPRNTSVANNMDRANSRSSKRRSKRRMNDRRSNS 542
QY 176 YKEIELEKEKDFENSCKESE-----PSQEOFPVEYEGH-----NNGL 214
DB 543 Y--TSRRDRERSYRNEEKREDDKPEVAPLVPEANWVPRFKSKTEKKLAPDGKTEL 600
QY 215 LKPRDE--EKDLSNK-----AMPSHCDYQONLGNBPDKYPCNGQYIPTFCDTSFTS 264
DB 601 LDKDEVERKMSLKLKLTLEMFDALISSEILAIANI-----SVWETNGETLKAVIEQIFLKA 656
QY 265 --RPEMOSYPTIYVYGPPLPSLNYHLNIORSGPPNPSNIFQADDSQIQNGIYVNNC 322
DB 657 CDEPHMSSYAOLOC--GKVVKELNDPIDETNEGTEGP-----KLVHLVAVARC 703
QY 323 HVNMWCMFTDQNNETD--CSENRSVHPSGNGCSMDRYVNSGFCVEYRERCMRDHGM 379
DB 704 HAE-----FDKG--WTDKLPTEDEGTPLPE--MSEYVAAASARRGL----- 744
QY 380 KHNQDREYVNOQFOEKLKILQ-KLTI-----LRLGLPES-GKTLVSRLLIGQNRDGI 432
DB 745 --GLVRFIGFLY--RLNLTGKMMFECEFRILMKDLTDSPEETLESVELLVNGEOLF 798
QY 433 STDQVFHHNDGYRYVNOLGDAHMDNQNRAKQALIDGSRPV-----IIDNTNIOAMKRP 487
DB 799 ETDSSF-----RTGGATLEGSOLLDSLGLILDNT-IQTAKISS 834
QY 488 YV-----EVAIGKGYVEFHE--PTWKFDEPELEKRNKHGVS--KKIQAOMDR 534
DB 835 RIKFKLIDIKELRHDKNNNSDKKDGPKTIQIHHEEENQRLKNNNSNSNSRTNNSNR 894
QY 535 YEQY-----MSISIVANSVEPSHKSTQRPBP 562
DB 895 HSFRRDAPAPSKDSFTTTRTYSQRSQRAAPRK 927

```

RESULT 14

```
YMF7_YEAST STANDARD; PRT; 1658 AA.
ID Q03661; Q04988;
AC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 187.1 kDa protein in G0A1-ERG8 intergenic region.
GN YMR219W OR YMR261.13 OR YMR959.01.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
[1]
RP SEQUENCE OF 1-711 FROM N.A.
RC STRAIN=S288C / AB972;
RA Dedman K., Brown D., Bowman S., Barrell B.G., Rajandream M.A.,
RA Walsh S.V.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 608-1648 FROM N.A.
RC STRAIN=S288C / AB972;
RA Skelton J., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RA Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; Z49809; CAA89934.1; -
DR EMBL; Z49939; CAA90190.1; -
DR SGD; S0004832; YMR219W.
KW Hypothetical protein.
SQ SEQUENCE 1658 AA; 187137 MW; 3893F968305A757D CRC64;

Query Match 3.6%; Score 117; DB 1; Length 1658;
Best Local Similarity 18.5%; Pred. No. 12;
Matches 106; Conservative 84; Mismatches 176; Indels 206; Gaps 28;

QY 9 KFLGPREVTSPRCKKLSTTESYVFHNHNSADFHRIQKTDNDW-----VPVTVIIDV 62
DB 322 KYMPRTDNTKPIVIEKYES-----DEHKVHQRYSEDGAFDFGVSNIISVDDE 368

QY 63 RGHSLQENKIKTTDLHRLHDEMPGNRPDVIESIDSVQLQEARPLIVSADDEIYTSKA 122
DB 369 SEDEESQAESYSANAENVYHHNEHELDKELTIEDIE-----SDSESQSAQES 416

QY 123 FIGP-----IVKPPKKRKNRGNRAHVLNGINDRGQKEKQKFNSEKSEIDNELFQFYKE 178
DB 417 EQGSEDDFEYKMKNEKSTSE---ETENTSESDQGFQAKDAYTKNVEQOENDE-----466

QY 179 IELEKEKGFENSEKESEPSQEQVFYFEGHNNGLLKPEDEKCOLSNKAMPCHSDYQON 238
DB 467 -----EPEKDDIIRSSLDKN-----FHGNNN---KSEYSENVLENETDPAIVE-REN 509

QY 239 LGNEPKYPCNGQVITFCDTSTFRPEWQSVYPIVPGPLPSLNLHLNQRFSGPP 298
DB 510 QINDVEGYDVGKSVES-----DLDEH-----531

QY 299 NPPSNIFQAQDSOIQNGYVNNCHVNNCMFTDONNEYTDGSENSSVHPS-----GNKC 354
DB 532 -SPDNLVLAARMLQ-----FQQRN-SNCPQKEQVESVILGHSNGS 573

QY 355 SMODRVVSGFCEVRRC-WKDHCMKNG--TDR-----FVNQOFQEEKLN-KLOK 402
DB 574 NLSGRSLD-----ESEQIPLKDTGTGNNNNLTKDRGLSSSVIEVEKYSKKLDGSTEK 629

QY 403 LLILLRGLPGSGKTTLSRILLQGNRQDGVFSDDDIFHHQDGYRYNNQLGDAHDNQNRA 462
```

```
Db 630 ELV-----PLSTDTTINSSLG-NEDSIYSLDD-----657
QY 463 KQAIQGRSPVVIDN-TNIOAWEMK--PYVEVAIGKY-----RVFHEP 504
Db 658 -----ADAISENLTDVPLMEIKTTPKYEVVISSEVSYSTYEDNTVAMPPOVEYTS 709
QY 505 ETWKKFDP-----EELEKRNKHGVSRKKTAQM 531
Db 710 ---FMNDPFLNDDYK--KHDLLKSTLAAL 736

RESULT 15
CN37_RAT STANDARD; PRT; 404 AA.
ID CN37_RAT
AC P13233;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 2',3'-cyclic nucleotide 3'-phosphodiesterase (EC 3.1.4.37) (CNP)
DE (CNPase).
GN CNP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87310616; PubMed=3040924;
RA Bernier L., Alvarez F., Norgard E.M., Raible D.W., Mentaberry A.,
RA Schreiber J.G., Sabatini D.D., Colman D.R.;
RT Molecular cloning of a 2',3'-cyclic nucleotide 3'-phosphodiesterase:
RT mRNAs with different 5' ends encode the same set of proteins in
RT nervous and lymphoid tissues.
RL J. Neurosci. 7:2703-2710(1987).
CC -!- CATALYTIC ACTIVITY: Nucleoside 2',3'-cyclic phosphate + H(2)O =
CC nucleoside 2'-phosphate.
CC -!- SUBCELLULAR LOCATION: FIRMLY BOUND TO MEMBRANE STRUCTURES OF BRAIN
CC WHITE MATTER.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M18630; AAA40939.1; -
DR PIR; A45670; A45670.
DR InterPro; IPR001230; Prenyl_site.
KW Hydrolase; Membrane; Brain
SQ SEQUENCE 404 AA; 45595 MW; 9C4CA888E6963BFF CRC64;

Query Match 3.6%; Score 116; DB 1; Length 404;
Best Local Similarity 24.7%; Pred. No. 2.2;
Matches 49; Conservative 32; Mismatches 65; Indels 52; Gaps 9;

QY 402 KLLILLRGLPGSGKTTLSRI-----LLQGNRQDGVFSDDDIFHHQDGYRYNV 448
DB 30 KTLFILRGLPGSGKSTLARNPWRSTTTAPRWCLLMLTRSFLLA-----72

QY 449 NOLGDAHDNQNRAKQAIQGRSP-----VIIDNTNIOAWEMKPYVEVAIGKYRVEPHE 503
DB 73 ---LGQTSPEYKRLDEDL-AGILPRDIRVLDVDDTNRERERLDQFEMADQYQVVLVE 129
QY 504 PETWKKFDEELEKRNKHGV-----SRKIAQMLDRYEYQMSIS--IVMNSVEPSHKSTQ 556
DB 130 PKTANRLDCAQLKKNQWOLLARIDDLKLLKPLEKDFLPLFGWFLTKKSETLRKDS- 188
QY 557 RPPPGQRQWCGSLGSH 574
DB 189 RP-----GSFQW--KLGNH 200
```

Tue Mar 11 10:11:24 2003

us-09-924-654-4.rsp

Page 11

Search completed: March 10, 2003, 17:53:28
Job time : 23 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 17:48:19 ; Search time 22 Seconds
(without alignments)
2547.562 Million cell updates/sec

Title: US-09-924-654-4

Perfect score: 3209

Sequence: 1 MSYGELEGKFLGPREVTSE.....RQRMGSLGSHNRVCVTNNH 583

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	195.5	6.1	232 2 F64576	hypothetical prote
2	164.5	5.1	162 2 T16163	hypothetical prote
3	161	5.0	149 2 D82178	hypothetical prote
4	154	4.8	400 1 ESB093	2',3'-cyclic-nucle
5	148.5	4.6	421 1 JCI317	2',3'-cyclic-nucle
6	147	4.6	420 1 ESM532	2',3'-cyclic-nucle
7	136.5	4.3	1551 2 F86342	hypothetical prote
8	136	4.2	944 2 T18637	hypothetical prote
9	134.5	4.2	2269 2 T18472	hypothetical prote
10	134	4.2	1271 2 A45552	glutamate rich pro
11	133	4.1	420 2 I56377	2',3'-cyclic-nucle
12	130.5	4.1	2485 1 H71621	serine/threonine-s
13	128	4.0	964 2 T21865	hypothetical prote
14	128	4.0	2500 2 G71609	hypothetical prote
15	128	4.0	2924 2 T18378	variant-specific s
16	127	4.0	275 1 TVRFR	transforming prote
17	127	4.0	1115 2 T41342	probable coiled-co
18	127	4.0	3119 2 T18414	protein g377 - mal
19	126	3.9	2269 2 T28677	thopyrin protein -
20	125	3.9	2523 2 T18477	hypothetical prote
21	123	3.8	1025 2 E86355	hypothetical prote
22	122.5	3.8	1804 2 H96597	hypothetical prote
23	121.5	3.8	2010 2 B71616	phosphatase (acid
24	120.5	3.8	871 2 D86355	protein T1615.12
25	120.5	3.8	1053 2 T51375	hypothetical prote
26	120.5	3.8	1059 1 A35210	ferroxidase (EC 1.
27	120	3.7	287 2 S45085	hypothetical prote
28	119	3.7	407 2 C45600	asparagine-rich bl
29	119	3.7	1790 2 S67593	transport protein

30	118.5	3.7	301 1 KIBPP4	polynucleotide kin
31	118.5	3.7	4981 2 T18489	hypothetical prote
32	118	3.7	969 2 T15446	hypothetical prote
33	117.5	3.7	952 2 S64473	translation initia
34	117.5	3.7	1844 2 D71612	hypothetical prote
35	117	3.6	773 2 F90537	lipoprotein (impor
36	117	3.6	1274 2 A89959	hypothetical prote
37	117	3.6	1658 2 S55101	hypothetical prote
38	116	3.6	404 2 A45670	2',3'-cyclic-nucle
39	115.5	3.6	1247 2 E71616	hypothetical prote
40	115	3.6	698 2 S49206	gl cyclin CLN1 - y
41	115	3.6	1164 2 S46769	hypothetical prote
42	115	3.6	1714 2 E71609	Ser/Thr protein ki
43	114.5	3.6	631 2 T32761	hypothetical prote
44	114	3.6	924 2 D81349	nitrate reductase
45	114	3.6	1044 2 S40704	hypothetical prote

ALIGNMENTS

RESULT 1
F64576
hypothetical protein HP0454 - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C>Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 24-Nov-1999
C:Accession: F64576
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKee
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watney,
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: F64576
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-232 <TOM>
A:Cross-references: GB:AE000560; GB:AE000511; NID:92313554; PIDN:AAD07520.1; PID:9231
C:Genetics:
A:Start codon: TTG
C:Superfamily: Helicobacter pylori hypothetical protein HP0454

Query Match 6.1% Score 195.5; DB 2; Length 232;
Best Local Similarity 32.0%; Pred. No. 1.1e-05;
Matches 57; Conservative 33; Mismatches 63; Indels 25; Gaps 8;

QY 394 EEKLNKLOKLLILRLGPGSGKTLRL-----LGNRGGIVFTSDDYRHOD--GYR 445
DB 8 QKRNNKSNKLVITINRAIPGGKSTLIRQIEELAKSLGHSIS--VHSIDEXFIQTDDEGIR 65

QY 446 YNV---NOLGDAHDMNQNRKAKIDQGRPYIINTNTIOAMEMKPYEVAIGKRYAEFH 502
DB 66 HYVVDKKKLNKYHONQDAEFKQALENRDIYVCNTNEMSQSKPYIDMAEFQKILLI 125

QY 503 EPETWKKFDEPELEKRNKKG--VSR--KKAQMDREYQNSISIVNNSVPSKSTQ 556
DB 126 D-----FKNRLETPMDYGDVQAQIKRPGIAKHVDYELRVL--VEPQDYERQ 175

RESULT 2
T16163
hypothetical protein F26A1.14 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16163
R:Fulton, L.
Submitted to the EMBL Data Library, May 1995
A:Description: The sequence of C. elegans cosmid F26A1.
A:Reference number: Z18469
A:Accession: T16163
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 1-162 <FUL>
A:Cross-references: EMBL:U27312; NID:g860679; PID:g860692; PIDN:AAA68256.1; CESP:F26A1.1
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:F26A1.14
A:Introns: 70/1; 125/2

Query Match 5.1%; Score 164.5; DB 2; Length 162;
Best Local Similarity 26.2%; Pred. No. 0.00071;
Matches 39; Conservative 35; Mismatches 46; Indels 29; Gaps 4;

Qy 394 BEKLNKLOKLL-----ILLRGLPGSGKTTLSRLLGNGRDGIVFSTDYFHHQDGYRY 446
Db 13 DAEISDIQKCLVEGHTIRILGVTSKSLARELVHSENGVIVKND-----VSN 64

Qy 447 NVNQLGDAHDWNNQAKQAIDQGRSPVVIDNTNIOAWEMKPYVEVAIGKGYRVEFHEPET 506
Db 65 NITR-----SVRFIDEDKHLIVDEENVQSSVKKFAELAVNGHVEIFVLEPDT 114

Qy 507 WNKFDPELEKRNKH-----GYSRKKIAQM 531
Db 115 DWRHDAIECKRSEKDEIGSIESKIMOL 143

RESULT 3
D82178
hypotheical protein VC1610 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
A:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: D82178
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qlin, H.; Dragol, I.; Sellers, B.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: R82035; MUID:20406833; PMID:10952301
A:Accession: D82178
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-149 <HEI>
A:Cross-references: GB:AB004239; GB:AB003852; NID:g9656130; PIDN:AAF94764.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1610
A:Map position: 1

Query Match 5.0%; Score 161; DB 2; Length 149;
Best Local Similarity 29.5%; Pred. No. 0.0011;
Matches 41; Conservative 26; Mismatches 46; Indels 26; Gaps 6;

Qy 404 LILLRGLPGSGKTTLSRLLGNGRDGIVFSTDYFHHQDGYRYNNVNLGDHWNQNR 462
Db 30 LTLIRGLPGSGKSTLAKTL-----SAVHLEADMYFVNPQGEYHFRPEYLAQAHEWCCOQT 84

Qy 463 KOAIDQGRSPVVIDNTNIOAWEMKPYVEVAIGKGYRVEF-----HEPETWKKFDPEELEK 517
Db 85 EYWLQGGKD-VVVSNTFVRHMEVAYRKLA--RQYRAKLTILVRC-----RY 129

Qy 518 RNKHGVSRRKKIAQMLDRYE 536
Db 130 QNVHGVDEATVERMRQWQ 148

RESULT 4
ESBOP3
2',3'-cyclic-nucleotide 3'-phosphodiesterase (EC 3.1.4.37) - bovine
N:Alternate names: CNPase; cyclic-CMP phosphodiesterase; nucleoside-2',3'-cyclic-phospha
C:Species: Bos primigenius taurus (cattle)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 11-Jun-1999
C:Accession: A26861; A27038; A29136; A43792
R:Vogel, U.S.; Thompson, R.J.
Nucleic Acids Res. 15, 7204, 1987
A:Title: Nucleotide sequence of bovine retina 2',3'-cyclic nucleotide 3'-phosphohydrolase

A:Reference number: A26861; MUID:88015580; PMID:2821502
A:Accession: A26861
A:Molecule type: mRNA
A:Residues: 1-400 <VOG>
A:Cross-references: GB:Y00405; NID:g261; PIDN:CAA68466.1; PID:g262
R:Vogel, U.S.; Thompson, R.J.
FEBS Lett. 218, 261-265, 1987
A:Title: Molecular cloning of the myelin specific enzyme 2',3'-cyclic-nucleotide 3'-p
A:Reference number: A27038; MUID:87247281; PMID:3036592
A:Accession: A27038
A:Molecule type: mRNA
A:Residues: 1-400 <VO2>
A:Cross-references: GB:M27606; NID:g162880; PIDN:AAA30457.1; PID:g162881
A:Experimental source: brain
R:Kurihara, T.; Fowler, A.V.; Takahashi, Y.
J. Biol. Chem. 262, 3256-3261, 1987
A:Title: cDNA cloning and amino acid sequence of bovine brain 2',3'-cyclic-nucleotide
A:Reference number: A29136; MUID:87137604; PMID:3029107
A:Accession: A29136
A:Molecule type: mRNA
A:Residues: 1-317,'A', 319-400 <KUR>
A:Cross-references: GB:J02659; NID:g162876; PIDN:AAA30456.1; PID:g162877
A:Note: this sequence has been corrected in reference A43792
R:Kurihara, T.; Fowler, A.V.; Takahashi, Y.
J. Biol. Chem. 262, 16754, 1987
A:Title: cDNA cloning and amino acid sequence of bovine brain 2',3'-cyclic-nucleotide
A:Reference number: A43792
A:Accession: A43792
A:Molecule type: mRNA
A:Residues: 316-320 <KU2>
A:Cross-references: GB:J02659
A:Note: this reference is a correction to reference A29136
C:Superfamily: 2',3'-cyclic-nucleotide 3'-phosphodiesterase
C:Keywords: phosphoric diester hydrolase

Query Match 4.8%; Score 154; DB 1; Length 400;
Best Local Similarity 28.6%; Pred. No. 0.012;
Matches 44; Conservative 37; Mismatches 49; Indels 24; Gaps 5;

Qy 394 BEKLNKLO--KLLILLRGLPGSGKTTLSRLLGNGRDGIVFSTDYFHHQDGYRYNNVNL 451
Db 20 EETVATLQECTLFLIRGLPGSGKSTLAFIVDYKRDGKMYVSADSYKITPGARGSFSE- 78

Qy 452 GDAHDWNNQAKQAIDQGRSP-----VIIDNTNIOAWEMKPYVEVAIGKGYRVEFHE 503
Db 79 ----EYKO-----LDEDLAACRRDRFVLYLDDTNHERERLEQLFELADQYQVQLVE 128

Qy 504 PETWKKFDPEELEKRNKHGVS--KKIAQMLDR 534
Db 129 PRTAWRLDCAQLKEKNQWLSADDLKLKPLGLEK 162

RESULT 5
JC1517
2',3'-cyclic-nucleotide 3'-phosphodiesterase (EC 3.1.4.37) long splice form - human
N:Alternate names: 2',3'-cyclic nucleotide-3'-phosphohydrolase; CNPase; cyclic-CMP ph
N:Contains: 2',3'-cyclic-nucleotide 3'-phosphodiesterase (EC 3.1.4.37) short splice f
C:Species: Homo sapiens (man)
C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 19-Jan-2001
C:Accession: JC1517; JC1518; A48934; A27703; PC4368
R:Monoh, K.; Kurihara, T.; Takahashi, Y.; Ichikawa, T.; Kumanishi, T.; Hayashi, S.; M
Gene 129, 297-301, 1993
A:Title: Structure, expression and chromosomal localization of the gene encoding huma
A:Reference number: JC1517; MUID:93314977; PMID:8392017
A:Accession: JC1517
A:Molecule type: DNA
A:Residues: 1-421 <MON>
A:Cross-references: DDBJ:D13144; DDBJ:D13145; DDBJ:D13146; NID:g219398; PIDN:BAA39694
A:Note: long form (isoform 1)
A:Accession: JC1518
A:Molecule type: DNA
A:Residues: 21-421 <MO2>

QY 61 DVGRHSYLO--ENKIKTTDLHRLHDEMPGNRPDVI-----ESTD-SQVLOEARPPV 110
Db 1208 ENNGKEVAKPTOKSOTTTSKXAVPDQP---PSIVTSLVSKKEETEKATPEEPPLK 1264
QY 111 SADDEIYSTSKAFIGIYKPPKRRNEG--RNEAHLNGINDRGQKQKFNSEKSEI 168
Db 1265 TKEEBEL-----IKKEEKRKQKAAKMQHRLBEIAKAEAMERKKKREKAKA 1315
QY 169 DNELOFYKEIELEKEKGFNSC-----KESEPQOFVPPYEGHNGL 214
Db 1316 -RAVLKAQEAEREKAYFCHSVIQIGRSVSLARTREAEKEGKGDFIIRDSN- 1373
QY 215 LKPDEKDLNSKAMPSCDYNLGNEDPKYPCNGOVIPTFCDDTSET 262
Db 1374 RPNDSRNCNRPKGDRNSKTRNGE-----STNEEISQTFITVSET 1417

RESULT 8
T18627
hypothetical protein B0001.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18627
R:Sims, M.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z18999
A:Accession: T18627
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-944 <WIL>
A:Cross-References: EMBL:Z69634; PIDN:CAA93451.1; GSPDB:GN00022; CESP:B0001.2
A:Experimental source: clone B0001
C:Genetics:
A:Gene: CESP:B0001.2
A:Map position: 4
A:Introns: 18/2; 75/3; 119/2; 350/3; 475/2; 487/3; 521/2; 601/3; 641/3; 749/2; 870/1

Query Match 4.2%; Score 136; DB 2; Length 944;
Best Local Similarity 20.7%; Pred. No. 0.56;
Matches 131; Conservative 83; Mismatches 205; Indels 214; Gaps 33;

QY 1 MSYGEIE-----GKFLGPREEVTSEPRCKKLKSTTE--SYVF-----HNHSNADF--- 43
Db 137 LSDGKIEVISKTAGVLTETKRNKSNKSEAKSEKYEYKIIGVYTGKQNRNNAVFVTCN 196
QY 44 -----HRIQKGTG---NDWPVTIIDVRGHSYLOENKIKTTDLHRLHDEMPGNRPDI 94
Db 197 LTYPGTDGIIKSQLQIGDWVEIRTKKDFVKYFSPNPTADTPRNVKHFKE-----I 249
QY 95 ESIDSOVLOEARPPLVSADDEIYSTSKAFIGIYKPPKRRNEGNEAHLV-NGINDR- 152
Db 250 EKIEI-----YSVTWAGN-----POVKIRNFOPFAGHVQGSVEDRF 287
QY 153 -GGQKEKQKFNSEKSEID-----NELFO-----FYKETELEKEKDGFNSECKESE 197
Db 288 LGTIAISKIVTSEKVDVLIKRLFOAGDGKQKATWTFKKNVKEKSESDSGSDSENENS 347
QY 198 PSQOEFVPPYEGHNGLLKPDEKDLNSKAMPSCDYNLGNEDPKYPCNGOVIPTFC 257
Db 348 SSTDS-----RPSRPKNVENKKS-----EEKINNSEK----- 376
QY 258 DTSFTSFRPEWQSVFVIVYGPPLPS-LNYHLNLTQRFSGPPN-----PPSIFQAQDS 311
Db 377 -----KEYSNVPKINYLHPPEFSPMNWQM-----GPPGIMHPRPIGHGAIPNP 422
QY 312 QIQNGYYVNVNCHVNCMTFDONNEYTDCSENRSVHPSGNCSQMDRVVSGFCEVRER 371
Db 423 LMYSPYMSMAIPQMM-----NRQPVPTNNSPEVHIKKG----- 458
QY 372 CWKDCMKDHKNGTDTRF-----VNOQFQEEKLNKLQKLLIL-LRGLPGSGKTTLSRLLGON 426
Db 459 -FKP-LISTENPLVFRGRCMCFSPESQKTNVLRKAVITSLK--PNNKGH-----LYGKE 509

QY 427 RDGIVFSTDDYFHHODGYRYNVNQ-LGDAHDWQNRKQAKIDQGRSPVLIIDNTNQAMEM 485
Db 510 KIAFWLWLDLDD--HKOSVYVYVSKNDGIEPGHFFNGLFA-----SNGDKWEC 552
QY 486 KPYVEVAIGK-----GYRVEFHEPETWKFDPPELEKRNK- 520
Db 553 KKYVK-PLGKLMGIVSNSIELQLIVETYPQSGERL--NPETYHSFIGIVDKFNKL 608
QY 521 -----HGYSRK-KIAOMLDRIYQMSISIVMNS 547
Db 609 PEDCSRGVYKISIKMNVNERNEWCWIVSKVFS 641

RESULT 9
T18472
hypothetical protein C0440c - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C:Accession: T18472
R:Lawson, D.; Bowman, S.; Barrell, B.
submitted to the EMBL Data Library, November 1998
A:Reference number: Z18937
A:Accession: T18472
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-2269 <LAW>
A:Cross-References: EMBL:AL008970; NID:el407852; PID:el332566; PIDN:CAA15615.1
C:Genetics:
A:Gene: C0440c
A:Map position: 3

Query Match 4.2%; Score 134.5; DB 2; Length 2269;
Best Local Similarity 19.3%; Pred. No. 2.3;
Matches 124; Conservative 82; Mismatches 210; Indels 225; Gaps 30;

QY 24 KKLKSTTESYFPHNSADFHRIQE-KTCNDWVPVTIIDVRGHSYLOENKIKTTDLHRLP 82
Db 9 KRRRYTYESINIINANKRYCMCNKMDND---NNIINERKY-FNNSSSIKNT----- 58
QY 83 HDEMPGNRPDVIESDSQVLOEARPPLVSADDEIYSTSKAFIGIYKPPKRRNEGNE 142
Db 59 -----GNNYK-NDNIDKSYDMCERSMMDRDEIYNN-----YK---NKNRNNYSNK 102
QY 143 AHVLANGIN-----DRGGQKE-----KQK- 161
Db 103 NNYNNINHMNDNRRTYKGDLDGRKQYLSNKHYESKNDNENSYNLKNYFNPIPKYNNN 162
QY 162 -----NSEKSEIDNELFOYKIEIELEKEK-----DGPNESC-----KESEPQEO 202
Db 163 NNNNNNNNDNRQNDNNIYKDYDDKNEKSKNTHSYGNFNNEFTNDPKQNTWSNES 222
QY 203 FVPFVEGH-----NNGLLKPDEKDLNSKAMPSCDYNLGNEDPKYPCNGQVI 253
Db 223 HMTODGHPYIYNNHNNNDGRK---NRKDISLSYGYNKISNDNLKENTSY----- 272
QY 254 PTFCDTSFTSFRPEWQSVFVIVYGPPLPSLNYHLNLTQRFSGPPNPSNIFQAQDSQI 313
Db 273 -----INQYNN-----SGGAN--RHVINAQHNQ 295
QY 314 QNGYYVNVNCHVNCMTFDONNEYTDCSENRSVHPSGNCSQMDRVVSGFCEVRER 373
Db 286 HNNNOHNN---NFNYHNYKNAPYQDKYKNGDSFKFSAYNABQHDHNMKNYQDTHKNNM 352
QY 374 KDHCMD-----KHNGT---DRFVNQOFQEEKLNKLQKLLILRLGLPGSGKTTLSRLLGON 426
Db 353 KHSTKGDTAISFNGQKVEGSKNKKQKERR-KPFK-----NKERGSSI-----N 395
QY 427 RDGIVFSTDDYFHHODG---YRYNVNQLGDA-----HDWQNRKQAKIDQ-GRSPVII 475
Db 396 RPYHNNNNNNNNNNNNNNNNNNHNFHDCDAERRNNKNNENWCEINSNELYEKDGMPYIN 455
QY 476 DNTNLOAMEMKPYVEVAIGKGYRVEFHEPETWKFDPPELEK-----RNKHGV----- 523

C:Accession: T18378
R:Baruch, D.I.; Pasloske, B.L.; Singh, H.B.; Bl, X.; Ma, X.C.; Feldman, M.; Taraschl, T
Cell 82, 77-87, 1995
A:Title: Cloning the P. falciparum gene encoding PfEMP1, a malarial variant antigen and
A:Reference number: 218925; MUID:95330812; PMID:7541722
A:Accession: T18378
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2924 <BAK>
A:Cross-references: EMBL:U27338; NID:9914918; PID:9914919; PIDN:AA60251.1
C:Genetics:
A:Gene: EMP1
A:Introns: 2476/3

Query Match	4.0%	Score 128	DB 2	Length 2924
Best Local Similarity	20.8%	Pred. No. 8.5		
Matches 88	Conservative 56	Mismatches 153	Indels 126	Gaps 22

```

QY 117 YGSKAFIPYKPPKKRRKRNGBAHVJLNGIN----- 150
Db 1472 YNLIKRIKIDPCT-----KKKODKTEHKCINGCINICECYRKYLEIKGNEMGINKKHNY 1526
QY 151 DRGQKQKOKFENSEKSEIDNELFOF-YKEIEE-LEKEKOGFENSEKSESEPSQOFPYFE 208
Db 1527 NSMDDKETIAYNVKSYFVDOGLEDTDYKKAQKAYDEK-----ERKKINGCT 1573
QY 209 GHNNGLLKDEDEKDKLSNKAAMPCHQYQNLGMEPRKYPONGOVIFPDCDTSTSPREN 268
Db 1574 GHDECKEKKERKKNITNLI--SELODKTISCONKHNPNK---TACD-PPSPPE- 1625
QY 269 QSVYPRIVEGPRPLPSLYNLIOFSGPPNP-----SNIFQADDSQIONGYVNN 321
Db 1626 -ETDPLDDTDPLDLD-DQTEQPKCPRPPTWTCVEKIAKELRYAEKIKINTELKNG 1683
QY 322 CHVNNCMCTFDONNETTDCSENSSVHPSPNGSGMDDRYVSNQFCVREKGCWMDHCMDKH 381
Db 1684 KDPNGKCNVKKKNGAVIGES-----CKFEQTY-ENSVNNNNNNKC-----KD 1725
QY 382 NGTDPR-VNOFOEKLNLKLOKLLIL-----LRLOPSSGKTYTL- 419
Db 1726 NQNERKIRGKMMFKYIGTIRKOLCIPRRREHMCDDLMSLGRRTTISDALLKIOEAA 1785
QY 420 -----RILLQNRDGIYFSTDDYFHQO--DGRYNNVQLOD---AHD-VNQRKAKQA 465
Db 1786 KSERDDIIRKLEON-----SCDE---HRICAMKYSFADLGIIRGRDLMMNNSKQK 1836
QY 466 IDQ 468
Db 1837 LOK 1839

```

Search completed: March 10, 2003, 17:54:55
Job time : 29 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 17:54:19 ; Search time 20 Seconds
(without alignments)
1229.260 Million cell updates/sec

Title: US-09-924-654-4

Perfect score: 3209
Sequence: 1 MSYGEIEGFLGPREVTSF.....ROMGSGLSHNRVCVTNNH 583

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 188354 seqs, 42170167 residues

Total number of hits satisfying chosen parameters: 188354

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3209	100.0	583	10	US-09-924-654-4
2	3205	99.9	594	10	US-09-925-300-1079
3	449.5	14.0	132	10	US-09-924-654-10
4	372	11.6	104	10	US-09-924-654-12
5	120	3.7	1111	10	US-09-815-242-12955
6	109.5	3.4	1338	10	US-09-402-100-4
7	108.5	3.4	291	10	US-09-925-300-968
8	108	3.4	665	9	US-09-820-843A-107
9	106	3.3	506	10	US-09-976-165-19
10	105	3.3	1093	10	US-09-801-368-392
11	104.5	3.3	2125	10	US-09-919-172-29
12	104	3.2	1435	9	US-10-153-273-4
13	104	3.2	1863	9	US-10-022-819-2
14	104	3.2	4636	10	US-09-835-996A-33
15	103	3.2	819	10	US-09-825-144-14
16	103	3.2	861	9	US-09-820-843A-109
17	102.5	3.2	1139	9	US-09-820-843A-15
18	102.5	3.2	1421	10	US-09-924-154-13
19	101	3.1	2507	9	US-09-819-104A-2

20	100	3.1	1295	10	US-09-726-949A-1	Sequence 1, Appl1
21	100	3.1	3594	9	US-10-150-821-4	Sequence 4, Appl1
22	100	3.1	3594	10	US-09-911-842-4	Sequence 4, Appl1
23	99.5	3.1	807	9	US-10-029-217A-2	Sequence 2, Appl1
24	99.5	3.1	935	9	US-10-029-217A-31	Sequence 31, Appl1
25	99.5	3.1	935	9	US-10-029-217A-32	Sequence 32, Appl1
26	99	3.1	340	9	US-09-789-054A-6	Sequence 6, Appl1
27	98.5	3.1	1719	9	US-10-024-450-4	Sequence 4, Appl1
28	98.5	3.1	3571	9	US-10-150-821-2	Sequence 2, Appl1
29	98.5	3.1	3571	10	US-09-911-842-2	Sequence 2, Appl1
30	98	3.1	532	10	US-09-891-160-2	Sequence 2, Appl1
31	98	3.1	782	9	US-09-935-868-48	Sequence 48, Appl1
32	98	3.1	782	9	US-09-935-868-52	Sequence 48, Appl1
33	97.5	3.0	1805	9	US-09-820-843A-73	Sequence 73, Appl1
34	97	3.0	370	9	US-10-086-823-8	Sequence 8, Appl1
35	97	3.0	370	9	US-10-139-583-37	Sequence 37, Appl1
36	97	3.0	370	9	US-10-039-847A-2	Sequence 2, Appl1
37	97	3.0	370	9	US-09-823-033-5	Sequence 5, Appl1
38	97	3.0	370	10	US-09-808-972-2	Sequence 2, Appl1
39	97	3.0	370	10	US-09-915-582-56	Sequence 56, Appl1
40	97	3.0	629	10	US-09-799-875-17	Sequence 17, Appl1
41	97	3.0	843	10	US-09-815-242-5370	Sequence 5370, Ap
42	97	3.0	843	10	US-09-815-242-12552	Sequence 12552, A
43	97	3.0	843	10	US-09-815-242-12898	Sequence 12898, A
44	97	3.0	1086	10	US-09-924-154-15	Sequence 15, Appl1
45	97	3.0	1863	9	US-09-734-672-2	Sequence 2, Appl1

ALIGNMENTS

```
RESULT 1
US-09-924-654-4
: Sequence 4, Application US/09924654
: Patent No. US20020146712A1
: GENERAL INFORMATION:
: APPLICANT: Good, Richard D.
: APPLICANT: Akedion, Ingrid E.
: APPLICANT: Sellhammer, Jeffrey J.
: APPLICANT: Coleman, Roger
: TITLE OF INVENTION: TUMOR SUPPRESSOR
: FILE REFERENCE: PC-0049 CIP
: CURRENT APPLICATION NUMBER: US/09/924, 654
: CURRENT FILING DATE: 2001-08-07
: NUMBER OF SEQ ID NOS: 21
: SOFTWARE: PERL Program
: SEQ ID NO 4
: LENGTH: 583
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No. US20020146712A1 496071CD1
US-09-924-654-4

Query Match      100.0%; Score 3209; DB 10; Length 583;
Best Local Similarity 100.0%; Pred. No. 9, 7e-247;
Matches 583; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSYGEIEGFLGPREVTSFPRCKKLTSTSYFHNHNSNADFRIOKGTGNDWPVTII 60
DB 1 MSYGEIEGFLGPREVTSFPRCKKLTSTSYFHNHNSNADFRIOKGTGNDWPVTII 60
QY 61 DVRGHSYLOENKIKTTDLHRLHDEMPGNRPDIESTDSOVLQEARPLVSADDEIVSTS 120
DB 61 DVRGHSYLOENKIKTTDLHRLHDEMPGNRPDIESTDSOVLQEARPLVSADDEIVSTS 120
QY 121 KAFIGPIYKPEKKRNKNGRNEAHVYNGINDRGCKEKQKNSKSELDNLFQFYKEIE 180
DB 121 KAFIGPIYKPEKKRNKNGRNEAHVYNGINDRGCKEKQKNSKSELDNLFQFYKEIE 180
QY 181 ELEKXDFENSKSESESOEFVPEYGHNGGLKPEEKDLSNKKAMPSEHCYOOVLG 240
DB 181 ELEKXDFENSKSESESOEFVPEYGHNGGLKPEEKDLSNKKAMPSEHCYOOVLG 240
```

Db 181 ELEKEKDFGFKSESEPSQEQVFPVEYEGHNNGLLKPDEEKKDLSNKAMPSCDQYQOQNLG 240
QY 241 NEPDYKPCNGQVPIPTCDTSFTSFRPEWQSVYFFIVPGPPLPSLNYHLNIQRFSGPPNP 300
Db 241 NEPDYKPCNGQVPIPTCDTSFTSFRPEWQSVYFFIVPGPPLPSLNYHLNIQRFSGPPNP 300
QY 301 PSNIFQAQDDSDIOQNGYVYNNCHVNNCMCTFDONNEYTDCSENRSVHPSPGNGCSWQDRY 360
Db 301 PSNIFQAQDDSDIOQNGYVYNNCHVNNCMCTFDONNEYTDCSENRSVHPSPGNGCSWQDRY 360
QY 361 VSGFCEVRERCWKDCHMDKHNGTDREVNQOQFQEEKLNKLOKLLILLRGLPGSGKTTLSR 420
Db 361 VSGFCEVRERCWKDCHMDKHNGTDREVNQOQFQEEKLNKLOKLLILLRGLPGSGKTTLSR 420
QY 421 ILLGNRQDGIIVSTDYDFHHQDGYRYNNVNLGDAHDWNNONRAKQAIQDGRSPVIIDNTNI 480
Db 421 ILLGNRQDGIIVSTDYDFHHQDGYRYNNVNLGDAHDWNNONRAKQAIQDGRSPVIIDNTNI 480
QY 481 QAWEMKPYVEVAIGKGYRVEFHEPETWKFDPPELEKRNKHGVSRRKKAQMLDRVEYQMS 540
Db 481 QAWEMKPYVEVAIGKGYRVEFHEPETWKFDPPELEKRNKHGVSRRKKAQMLDRVEYQMS 540
QY 541 ISIVNSVEPSHKSTQRPPOGQORWGSGSLGSHNRVCVTNNH 583
Db 541 ISIVNSVEPSHKSTQRPPOGQORWGSGSLGSHNRVCVTNNH 583

RESULT 2

US-09-925-300-1079
; Sequence 1079, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCS/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1079
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (430)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1079

Query Match 99.9%; Score 3205; DB 10; Length 594;
Best Local Similarity 99.8%; Pred. No. 2.1e-246;
Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSYGEIEGKFLGPREEVTSERPCCKLKSTTSYVFNHNSNADFHRIQEKTDGNDWVPVTII 60
Db 12 MSYGEIEGKFLGPREEVTSERPCCKLKSTTSYVFNHNSNADFHRIQEKTDGNDWVPVTII 71
QY 61 DVRGHSYLOENKIKTTDLHRPLHDEMPGNRPDVIIESDSQVLQEARPLVSADDEIYSTS 120
Db 72 DVRGHSYLOENKIKTTDLHRPLHDEMPGNRPDVIIESDSQVLQEARPLVSADDEIYSTS 131
QY 121 RAFIGIYKPPPEKKRNGRNEAHVNLGINDRGQKQKQKENSEKSETDNLNLFQYKIE 180
Db 132 RAFIGIYKPPPEKKRNGRNEAHVNLGINDRGQKQKQKENSEKSETDNLNLFQYKIE 191
QY 181 ELEKEKDFGFKSESEPSQEQVFPVEYEGHNNGLLKPDEEKKDLSNKAMPSCDQYQOQNLG 240
Db 192 ELEKEKDFGFKSESEPSQEQVFPVEYEGHNNGLLKPDEEKKDLSNKAMPSCDQYQOQNLG 251

QY 241 NEPDYKPCNGQVPIPTCDTSFTSFRPEWQSVYFFIVPGPPLPSLNYHLNIQRFSGPPNP 300
Db 252 NEPDYKPCNGQVPIPTCDTSFTSFRPEWQSVYFFIVPGPPLPSLNYHLNIQRFSGPPNP 311
QY 301 PSNIFQAQDDSDIOQNGYVYNNCHVNNCMCTFDONNEYTDCSENRSVHPSPGNGCSWQDRY 360
Db 312 PSNIFQAQDDSDIOQNGYVYNNCHVNNCMCTFDONNEYTDCSENRSVHPSPGNGCSWQDRY 371
QY 361 VSGFCEVRERCWKDCHMDKHNGTDREVNQOQFQEEKLNKLOKLLILLRGLPGSGKTTLSR 420
Db 372 VSGFCEVRERCWKDCHMDKHNGTDREVNQOQFQEEKLNKLOKLLILLRGLPGSGKTTLSR 431
QY 421 ILLGNRQDGIIVSTDYDFHHQDGYRYNNVNLGDAHDWNNONRAKQAIQDGRSPVIIDNTNI 480
Db 432 ILLGNRQDGIIVSTDYDFHHQDGYRYNNVNLGDAHDWNNONRAKQAIQDGRSPVIIDNTNI 491
QY 481 QAWEMKPYVEVAIGKGYRVEFHEPETWKFDPPELEKRNKHGVSRRKKAQMLDRVEYQMS 540
Db 492 QAWEMKPYVEVAIGKGYRVEFHEPETWKFDPPELEKRNKHGVSRRKKAQMLDRVEYQMS 551
QY 541 ISIVNSVEPSHKSTQRPPOGQORWGSGSLGSHNRVCVTNNH 583
Db 552 ISIVNSVEPSHKSTQRPPOGQORWGSGSLGSHNRVCVTNNH 594

RESULT 3

US-09-924-654-10
; Sequence 10, Application US/09924654
; Patent No. US20020146712A1
; GENERAL INFORMATION:
; APPLICANT: Goold, Richard D.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Coleman, Roger
; TITLE OF INVENTION: TUMOR SUPPRESSOR
; FILE REFERENCE: PC-0049 CIP
; CURRENT APPLICATION NUMBER: US/09/924,654
; CURRENT FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID NO. US20020146712A1 g533948
US-09-924-654-10

Query Match 14.0%; Score 449.5; DB 10; Length 132;
Best Local Similarity 82.1%; Pred. No. 1.4e-28;
Matches 87; Conservative 0; Mismatches 0; Indels 19; Gaps 1;

QY 1 MSYGEIEGKFLGPREEVTSERPCCKLKSTTSYVFNHNSNADFHRIQEKTDGNDWVPVTII 60
Db 46 MSYGEIEGKFLGPREEVTSERPCCKLKSTTSYVFNHNSNADFHRIQEKTDGNDWVPVTII 86
QY 61 DVRGHSYLOENKIKTTDLHRPLHDEMPGNRPDVIIESDSQVLQEAR 106
Db 87 DVRGHSYLOENKIKTTDLHRPLHDEMPGNRPDVIIESDSQVLQEAR 132

RESULT 4

US-09-924-654-12
; Sequence 12, Application US/09924654
; Patent No. US20020146712A1
; GENERAL INFORMATION:
; APPLICANT: Goold, Richard D.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Coleman, Roger
; TITLE OF INVENTION: TUMOR SUPPRESSOR

```

; FILE REFERENCE: PC-0049 CIP
; CURRENT APPLICATION NUMBER: US/09/924,654
; CURRENT FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PERL Program
; SEQ ID NO 12
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020146712A1 PANC 1B
; NAME/KEY: unsure
; LOCATION: 5,44, 75, 101
; OTHER INFORMATION: unknown or other
; US-09-924-654-12

```

```

Query Match          11.6% Score 372; DB 10; Length 104;
Best Local Similarity 94.6%; Pred No. 1.4e-22;
Matches 70; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 1 MSYGEIEGFLGPREVESEPRCKKLKSTESYVFHNSNADFHRIQEKTGNDVPEVTII 60
    |||||
DB 31 MSYGEIEGFLGPREVESEPRCKKLKSTESYVFHNSNADFHRIQEKTGNDVPEVTII 90
    |||||
QY 61 DVRGHSYLOENKIK 74
    |||||
DB 91 DVRGHSYLOENKIK 104

```

```

RESULT 5
US-09-815-242-12955
; Sequence 12955, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12955
; LENGTH: 1111
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-09-815-242-12955

```

```

Query Match          3.7% Score 120; DB 10; Length 1111;
Best Local Similarity 18.1%; Pred. No. 0.38;
Matches 117; Conservative 69; Mismatches 224; Indels 238; Gaps 25;

```

```

QY 3 YGEIEGKFLGPR-----EVTSEPRCKKLKSTESYVFH-----NHSNADFHRIQEKTN 52
    |||
DB 47 YSRPRGKFRPPMSVAYENENEQASDTISDEKQY--HRDYKQSHDSRSQKRRRRRNQ 104
    |||
QY 53 DWVPVTTIDVRGHSYLOENKIKTTDLHRLPHDEMPGNRPVIESISQVLO-----103
    |||
DB 105 TTEQNYSEQRGNSKISQOSIKKD-HSHYHTKPPOTYSALNGIEKTHKPTHMYSN 163
    |||
QY 104 ---EAPPLVSADDEIYSTSK---AFIG-----PIYKPEKKRRNEGNEAH 144
    |||
DB 164 NTNHRAKDSTPDYHKSFTSEVPSAIFGTMKPKLENGRIPVSKSEKVESDKQYDY 223
    |||
QY 145 VLVGINDRGQKQKQKFN-----EKSEIDNEL-----F 173
    |||
DB 224 VAKTQTSQNKQLEQEKQNDQSVYKQGTASKSDENASTTKSMNPYKSVNTIKIENIYAS 283
    |||
QY 174 OFYKET-----ELEKEKDFGNSCKSESPSO-BOFVPFYE 209
    |||
DB 284 QIYEELRERERKVLQKRFKALQOKREHKNEBQDAIQALDEMYAKQAEKYV----- 338
    |||
QY 210 HNNGLKPPDEKKDLNKAHPSHCDYQON--LQNEPDKYPC--NGOVIPTFCDTSFSPR 265
    |||
DB 339 -GDSLNDSDLDLNDSTASQHTNGIENETVSNDEKQASIQNEDTNDTHDES----- 392
    |||
QY 266 PEMQSVYPTVYGPPLPSLNHLNIOFSGPPNPSPNIFQADDSQIQNGIYVNNCHVN 325
    |||
DB 393 -----PYNEEVSINQVSTTKQJSDDEVTVSNV-TSQHQSALQHNVEVN----- 435
    |||
QY 326 WNCMTFDQNNETDCSENSSVHPSGNGCSMDRIVSYNGCFEVRKNCMDHCKDKNGTD 385
    |||
DB 436 -----DK-----D 438
    |||
QY 386 RFVNOQFOEEKLNKQLKLLKGLPGSGTTLRIILGQNRDGIYSTDDYFHHODYR 445
    |||
DB 439 ELKNO-----SRLIADSEEGATNKEL-----460
    |||
QY 446 YVNVQIGDAHDMNQNRKAQIDQGRSPVITDNTNIOAMENKPYEVAIGKGYVEHEPE 505
    |||
DB 461 YSGSQIDDAEFYELNDTE--VDEDITSNIEDNTNRNASEH--HVDAPKQGEYAVTESQVN 516
    |||
QY 506 TMMKFPDELEKRNKHGVSRRKIAQMLDRKEYQMSI-SIVMNSVEESH 552
    |||
DB 517 NDKTVDNELIAPRH-----KKDDQ-----NLVSNLSKTNDVNDNH 554
    |||
RESULT 6
US-09-402-100-4
; Sequence 4, Application US/09402100
; Patent No. US20010019834A1
; GENERAL INFORMATION:
; APPLICANT: Daewoong Pharmaceutical Co, LTD
; APPLICANT: Kim, Byung-O
; APPLICANT: Shin, Sung-Seup
; APPLICANT: Yu, Young-Hyo
; APPLICANT: Park, Myung-Hwan
; APPLICANT: Choi, Deok-Joon
; APPLICANT: Jung, Hyung-Jin
; TITLE OF INVENTION: Recombinant Microorganisms Expressing Antigenic Proteins of Hel
; FILE REFERENCE: 0136/05140
; CURRENT APPLICATION NUMBER: US/09/402,100
; CURRENT FILING DATE: 1999-09-27
; EARLIER APPLICATION NUMBER: KR 97-11950
; EARLIER FILING DATE: 1997-03-31
; EARLIER APPLICATION NUMBER: KR 97-11951
; EARLIER FILING DATE: 1997-03-31
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1338
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

```

US-09-402-100-4
OTHER INFORMATION: CagA/CTXA2B Chimeric protein

Query Match	3.4%	Score 109.5	DB 10	Length 1338
Best Local Similarity	19.1%	Pred. No. 3.4		
Matches 128	Conservative	98	Mismatches 240	Indels 205
			Gaps	33

QY 3 YGIECKFLGPBREVTSERPCCKKLKSTTSSYVFHNHNSADFHRIOEKTGNWDWPVTIIDV 62
 ||: ||: ||: ||: | :|::||: | | ||
Db 473 YGKKADKALDREKNWTLQOSLKH-----DGMFVDYSNEFYTN-ASKNPNGVGVT---- 522

QY	63	RHGSYLQE--NKIK-----TTDLRPLHDMPGPNPDVIESIDSVLQEARPP	108
		: :	
		: :	
Dd	523	NGVSHLEVGFNKVAIFNLPDLNNLAITSFVRNLDEKL-----TTGLSPQEAANKLIKLD	576

```

109 QY LVSADDEIYSTKAF-----IGPIVKPEKKKNEGRNEAHVYNGINDRG 153
      : : : : :
577 Db FLSSNKLGVKTLNFKAVADKNTGNYDEVKKAQDLKSLRKREHLEKEVEKKEKLSK 636
      : : : : :

```

154 GQKEQO-----KFNSEKSEI-----DNLDFYKIEIELEKE-KDGFNSCKSESEP 198
| | | : | | |
637 GHKNKEAAQAQNSQKDEIFALINKEANDRAIIYAQNKLGIKRELSDKLEWNNKKLD 696

QY	199	SOEQFVPFYEGHNNGLLKDPDEKKDLNKA	KAMP	SHCDYQONLGNPEPKYPCNGQVIT	PCD	258
		:	:	:	:	
Db	697	FKSTDFNGKNGKDFSKAEETLKALG	SV	-----	KDLG	730

QY	259	TSTSTFRPEWQSVYFIVPGPPLSLNHLNIQFSGPPNP	SNITQAQDDSOIQNGY	317
Db	731	-----INPEWIS-----	-----KVENLNAALN-----	770
			-----EFKINGKNKDFSKVYQAKSD-----	

```

318 YVNNCHVNNCMTFDQNEEYTDCESENRSVSHPGNGCSDMODRYSVGFECEVRERCWKDHC 377
      + : + : + : + : + : + : + : + : + : + : + : + : + : + : + : + :
771 *VDVII-----NQKVTDKVLNQLNO-----AVSVAKAGTDFSRVEQA-----L 807

```

378 MDKHGTDREVNQQTQE-EKLNKLOKLLIL-----LRG-LPGSGKTTLSRLLGNRDG 429
 808 ADLKFNSEQLAAQQAQKNESLNARKKSEIYQSVKGVNGTIVGSLGSAQAEITLSKNFSD 867

QY 430 IYFSTDDYFHODGYRYNVN-OLGDAHWNQNRKQAIDQGRSPVIIDNTNIQA-----W 483
 868 I-----KKELNAKLGFNNNNNGLK-----NEPYAKVNNKKAGQAASL 907
 Db

484 QY ENKPIVEVA-----IGK--GYRVFHFEPETWVKDPPELEKRKHG 522
- : || : | : : |
908 Db EEPPIAQVAKKNVNAKIDRLNQIATSLGVVGQAGPLRKH-----DKVDLSKVG 957

523 VSR-KKIAQMLDRYEQMS-----ISTVMNSVPSHKSTORPPP---PGRQRWGSGLG 572
 ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
958 LSRNQLAQKIDNLNQAYSEAKGFFGNLEQTDLKDSTKHNPNNLWESAKKVPA^{SLS} 1017

Ddb

QY 573 SHNRVCVTNNH 583
:
||:
Db 1018 AKLDNYATNSH 1028

RESULT 7
US-09-925-300-968

Sequence 968, Application US/09925300
Patent No. US20020151681A1
GENERAL INFORMATION:
APPLICANT: Craig Rosen,
INVENTOR:

APPLICANT: Steve Ruben
 TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 FILE REFERENCE: PA101
 CURRENT APPLICATION NUMBER: US/09/925,300

CURRENT FILING DATE: 2001-08-10
 PRIOR APPLICATION NUMBER: PCT/US00/05988
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/124,270

PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1890
SOFTWARE: PatentIn Ver. 2.0

```

; SEQ ID NO 968
; LENGTH: 291
; TYPE: PRT
; ORGANISM: H
US-09-925-300-9

```

Query Match 3.4%; Score 108.5; DB 10; Length 291;
Best Local Similarity 23.9%; Pred. No. 0.49;
Matches 57; Conservative 24; Mismatches 78; Indels 79; Gaps 12;

QY 87 PGNRPVIESIDSQLQEARPLVSADDEIYSTSKAFIGIYXP-----PEKKRNEGRN 141
| | | | : | | : | | : | |
Db 96 PQPRGVIRAL-----GPP-----GVRRRCQGISPEEEERRVR 132

```

QY      142  EAHVLGINDRGQKQKQF-NSEKSEIDNELFQFYKEITELEKXGDFE-----NSC 193
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      133  ERNKAAAKCNRKRKELTDFLQATOKLDEKSGLORETELQKQKLERLEVLIAHRPIC 192

```

QY 194 KESEPQEQFVPEYEGHNNGLLPDEKKDLSNKAMPSHC DYQQNLGNEDPKYC-----N 249
| | : : | : | : |
Db 193 KIPEGAKEGDT---GSGTSSP-----PAPC-----RPVQLSLSP 226

```

QY 250 GQVI-----PTFCDT-SFTSRPQSWYPIVPGPLPSLVHLNIQRESGPPN 299
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 227 GPVLEPEALHTPTLMTTPTSLTPTPSLVFTYP-----STPEPCASAHRRKSSSSGDPDS 279

```

RESULT 8
US-09-820-843A-107
: Sequence 107. Application: US/09820843A.

```

; sequence 1U/, Application US/09820843A
; Publication No. US20030039963A1
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE DETERMINATION OF C

```

```

; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CA
; TITLE OF INVENTOR: USEFUL AS ANTI-INFECTIVES
; FILE REFERENCE: Q63915
; CURRENT APPLICATION NUMBER: US/09/820,843A
; CURRENT FILING DATE: 2001-03-30

```

```

; CURRENT FILING DATE: 2001-03-30
;
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 107
; LENGTH: 665
;

```

```

; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: misc feature

```

```

; NAME/KEY: misc.feature
; OTHER INFORMATION: hypothetical protein
; NAME/KEY: misc.feature
; OTHER INFORMATION: g113845248
US-09-820-843A-107

```

Query Match 3.4%; Score 108; DB 9; Length 665;
Best Local Similarity 19.7%; Pred. No. 1.7;
Matches 68; Conservative 55; Mismatches 117; Models 106; Gaps

	mismatches	68;	Conservative	33;	Mismatches	117;	Indels	106;	Gaps	
QY	2	SXGEI-EGKFLGPREVTSEPRCKLKSTES-----YVFHNHNSADPHR-----	45							
		: : : :		:	:	:	:	:	:	
Dd	28	SYNDVIDENIKKEKSSEPFIPIKNKDIVHTNNNOYNIHNKSNKTHICYGTHTS	A7							

28	SINDAYEDINAREENSSSEPP	IFIKNKIDNVHTRNNQYNLHANKSNKTHLTGTHTS	87
QY	46	-IQEKTGNDWVPVTIDV-----RGHSVLOE-NKTKTDLHR-----	80
		: : : : : : : : : : : : : : : : :	
Db	88	FLQNTICDQVDVNDKSDINITKEKDNNNNCKTKOIEKKNKTKSDHONELNLOS	147

81	QY	-----PLHDEMPGNPDVIESIDSQVLQEARPPVSADDETSYSKAFIG	125
06	PQNCIINDCVDVDNNKDSINNIITAEKDUNNNGGTRQIEEKNINKSDJHRQNELNEQS		147
148	Dd	GNEODINKKEGGKODISNAENCKDVKVEGY-KELFKKKPEFTSDOHKVFENKS---	203

```

126 PIYKPEKKRNE-----GRNEAHVNGI-----NDRGQKEKQKFENSEIDNE 171
140 GANEQDINNAGKQDISNSNENADNDKGV-KELEKKKEEKISDDHKRVENKKS--- 203
204 DDHKVEENKSDDHKVEENKSDDHKIEVKVEPHEDEDEDEKKEKSEKNNENKQK 263

```

172 LQFYKEI-----EELEKEKDGSENSCKES-----EPSQEQFV--PFYEGHNGL 214

Db 264 NDEONDEI SDEDEYDDVDEEBKKNNDIDDDKKTDTXTHLEEEENEIIEKFSDDKKNKG 323

QY 215 LK-----PDEKKLSKKAMP SHDYQONLGNEDK 245.

Db 324 NKDKKEKSKDTEKEKSKDIEKEKSKDKKEKSKDKKEKSKDKKEK 369

RESULT 9

```

US-09-976-165-19
: Sequence 19, Application US/09976165
: Patent No. US20020107383A1
GENERAL INFORMATION:
APPLICANT: FUJIMURA, TSUTOMU
APPLICANT: MATANABE, TAKESHI
APPLICANT: HORIE, MASATO
TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
TITLE OF INVENTION: SKELETAL MUSCLE-SPECIFIC UBQUITIN-CONJUGATING ENZYME
FILE REFERENCE: Q-53599
CURRENT APPLICATION NUMBER: US/09/976,165
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/565,538
PRIOR FILING DATE: 2000-05-05
PRIOR APPLICATION NUMBER: 09/055,699
PRIOR FILING DATE: 1998-04-07
PRIOR APPLICATION NUMBER: 08/820,170
PRIOR FILING DATE: 1997-03-19
PRIOR APPLICATION NUMBER: JP 63410/1996
PRIOR FILING DATE: 1996-03-19
PRIOR APPLICATION NUMBER: JP 6163/1997
PRIOR FILING DATE: 1997-03-05
NUMBER OF SEQ ID NOS: 95
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 19
LENGTH: 506
TYPE: PRT
ORGANISM: Homo sapiens
US-09-976-165-19

```

Query Match	3.3%	Score 106	DB 10	Length 506
Best Local Similarity	19.9%	Pred. No. 1.7		
Matches 68; Conservative	51;	Mismatches 135;	Indels 88;	Gaps 14

```

OY 74 KTTDHLRHLHEMGNRDVI-----ESIDSVLQEARPLASADELSTSKAFIGYK 139
Db 135 KYAELNKLTYD-----RRQIINALEYPTBEECEMNSDEBSDEEVQDTPMSMPL-- 188
OY 130 PPEKKARNEGRNEAHLVINGIDRGQCKEYKQFNSKSEIDNELFOYKTEIBLEKEKDG 188
Db 189 -----EGGEENEPKNEPEVKAKEKVPK-----EIEPVKDEKEVKAKEIPEVAAEKAD 237
OY 190 ENSCKESEPSOQO-----FVPEYEGANNGLKJDEEK-----KDISNKAMPSCDYQ 237
Db 238 SKDCHEARPEVKEDEPKVEPYQKADKDQPKATEKAKARAARETHKRYRPEBLRNSVDLKR 297
OY 238 NLGNPEDKYPCNGOVIYIPFCCTSTSTSRPEKQSYPRFIVYGPPL-----PSLNYHLNIO 292
Db 298 ARKQKPKKEDBKG--IPDY-----WLIYLVKNVDKLGPMIOKYDEPLKFLSDVS 344
OY 293 -RFSGPMPNPSPINIFOADSDSIOJNGYVYNCHV-----MNCCTFPODNEYTDCSE 342
Db 345 LKFSRQGPVASTPEFH--FLPRPYFNEVLYVTYIIKAKRPHNDPFFSNGWGEDIC-- 399
OY 343 NRSSVHPSGNGCSM-----QDKYVSGFCEVER 371
Db 400 -----KCKIDRRGRGDVYTTTQSRTTAAEGEIIQOR 432

```

RESULT 10
US-09-801-368-392
; Sequence 332, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert

```

1  APPLICANT: Call, Brian
2  APPLICANT: Hecht, Peter
3  APPLICANT: Holtzman, Doug
4  APPLICANT: Madden, Kevin
5  APPLICANT: Maxon, Mary
6  APPLICANT: Milne, Todd
7  APPLICANT: No. us20020128250Alman, Thea
8  APPLICANT: Royer, John
9  APPLICANT: Salama, Sofie
10 APPLICANT: Sherman, Amir
11 APPLICANT: Silva, Jeff
12 APPLICANT: Summers, Eric
13 TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
14 FILE REFERENCE: 109272.147
15 CURRENT APPLICATION NUMBER: US/09/801,368
16 CURRENT FILING DATE: 2001-03-07
17 PRIOR APPLICATION NUMBER: US 09/487,558
18 PRIOR FILING DATE: 2000-01-19
19 PRIOR APPLICATION NUMBER: US 60/160,587
20 PRIOR FILING DATE: 1999-10-20
21 NUMBER OF SEQ ID NOS: 440
22 SOFTWARE: PatentIn version 3.0
23 SEQ ID NO 392
24 LENGTH: 1093
25 TYPE: prt
26 ORGANISM: Saccharomyces cerevisiae
27 US-09-801-368-392

```

Query Match	3.3%	Score 105;	DB 10;	Length 1093;
Best Local Similarity	17.3%	Pred. No. 5.8;		
Matches 97; Conservative	79;	Mismatches 227;	Indels 158;	Gaps 18;

QY	46	IOEKGNMVAWPT--IIDVRGSHYLOENKIKTTDDLRHLDMPG-----	88
Db	60	VMARKDWMINTIYOFKIAOFSKTRKRIKILEKESMDIOHEVUGGYGRFOGTWPLRDSAK	119
QY	89	---NRPDVIESIDSOVL-----QEARPLVSADEIYTSKA--FIGPIYKPEKKRNE	138
Db	120	FLVNYKEIIDIYVNSILFFQFDPPNPPRKBSKNSILKRTSPGKITSPSSYNKTPRKNS	179
QY	139	GRNEAHVLGINDRGOKOKOKFNS-EKSEIDNELFOYKIEIELEKOKGFENSCKESE	197
Db	180	SSSTATTATTAANKG--KKNASINQPNPSPLONLVFO-----T	215
QY	198	PQOQDFYFVEEGHNGLLKPBEKKDJSNKAMPSSHOCYQONLGNRPDKYPCONGVI---P	254
Db	216	PQOQVYVNSSMNIMN-----NDNHTTFNFPNDTRHLLINNISNSMOSTIIOOK	265
QY	255	TFCDFSTFSFPEMOSVYPIYVGPRLPSLNYHLNIDORSGBPSPNSITFOAODDSOIQ	314
Db	266	SIHEMSFNNNSYATOKPQLOFF-----PPTMLQKNV-ALNPNPNNDNSYSVSHIDVIN	319
QY	315	NGYVNNCHVMNCKTFD-----QNNETLYDCSENNBSVYPSGNGCSMODRIVSNG	364
Db	320	SSNNNNNGNNNLILVPRGPMOSQOOQHHELYLTNNEHSMOSITNGSKRRKRLN-	378
QY	365	PCFYARECMKOCMKHNGJTDIFYVQOFOEKLKJLKLILLRGPSGKTLIRLLG	424
Db	379	-----QSNEQOFIYOQEKIORHFKKOPBLMOSF-----	408
QY	425	QNRDQIVFSTDDYFHNDQGYRYNQLDADADWMONRAKQAIIDQGRSPVILIDNTNIQAME	484
Db	409	QNPND-----NH-----NEYDSSNGNNNNNTVASNGSSIEVPSNNDSMN	455
QY	485	MKPYEVALIGKYRVEFHEPETWKKFDEBELKRRKHGYSKKIAQMLDRLEYOMSSIV	544
Db	452	MSRSMPFESAG-----NTSSONKLENNMTDQYKOTILTI	487
QY	545	MNS-----VPSHKSTQRRP	560
Db	488	ISSERSSDVDQALLATLYPAR	508


```

OY      357 QDRYVSN 364
Db      1264 QDRGNSG 1271

RESULT 13
US-10-022-819-2
: Sequence 2, Application US/10022819
: Publication NO. US20030027166A1
: GENERAL INFORMATION:
:   APPLICANT: ALLEN, Antonette C. P.
:             OLSEN, Sheri J.
:             LAWRENCE, Tammy
:             ANGELLY, Tracy S.
:             RABIN, Mark B.
: TITLE OF INVENTION: CODING SEQUENCE HAPLOTYPE OF THE HUMAN
:             BRCA1 GENE
: NUMBER OF SEQUENCES: 67
: CORRESPONDENCE ADDRESS:
:   ADDRESSEE: Morgan Lewis & Bockius LLP
:   STREET: 1111 Pennsylvania Avenue
:   CITY: Washington DC
:   STATE: District of Columbia
:   COUNTRY: USA
:   ZIP: 20004
: COMPUTER READABLE FORM:
:   MEDIUM TYPE: Floppy disk
:   COMPUTER: IBM PC compatible
:   OPERATING SYSTEM: PC-DOS/MS-DOS
:   SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: US/10/022,819
:   FILING DATE: 22-Apr-2002
:   CLASSIFICATION: <Unknown>
:   PRIOR APPLICATION DATA:
:     APPLICATION NUMBER: 09/074,452
:     FILING DATE: 1998-05-06
: ATTORNEY/AGENT INFORMATION:
:   NAME: <Unknown>
:   REGISTRATION NUMBER: <Unknown>
:   REFERENCE/DOCKET NUMBER: 044921-5049-01-US
: TELECOMMUNICATION INFORMATION:
:   TELEPHONE: 202-739-3000
:   TELEFAX: 202-739-3001
: INFORMATION FOR SEQ ID NO: 2:
:   SEQUENCE CHARACTERISTICS:
:     LENGTH: 1863 amino acids
:     TYPE: amino acid
:     STRANDEDNESS: single
:     TOPOLOGY: linear
:   MOLECULE TYPE: protein
:   HYPOTHEICAL: NO
:   ANTI-SENSE: NO
:   FRAGMENT TYPE: N-terminal
:   ORIGINAL SOURCE:
:     ORGANISM: HOMO SAPIENS
:     STRAIN: BRCA1
:     HAPLOTYPE: OM14
:   POSITION IN GENOME:
:     CHROMOSOME/SEGMENT: 17
:     MAP POSITION: 17q21
:   SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-022-819-2

Query Match      3.2%; Score 104; DB 9; Length 1863;
Best Local Similarity 18.7%; Pred. No. 15;
Matches 94; Conservative 67; Mismatches 158; Indels 184; Gaps 24;
OY      65 HSYLOENIKTTDLHRPLHDEMPGNRPDVI---SIDSQVLOE-----ARP 107

```

[illegible]

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-835-996A-33

Query Match 3.2%; Score 104; DB 10; Length 4636;

Best Local Similarity 23.6%; Pred. No. 51;
Matches 50; Conservative 27; Mismatches 79; Indels 56; Gaps 13;

QY 187 DGFENCKESESQEFVPEYEGHNNGLKPD-----EKKDLSNKAMP--CDYQQ 237
DB 3414 DG-ENDC-----GDSDELNCDFHVLCSGQFKCTKNQKICIPVNLRCNGQD 3457
QY 238 NLGNEDKYPK-NGQVPTF--CDTSFTSRPEWQSVYFVYGGPPLSLNHLNIQRF 294
DB 3458 DCGDEDEDCFENSCSPDYFOCKTKHCISKLW-----VCDEDPDCADASDEANCDKK 3511
QY 295 SGPPNPSPSIFQAQDSQIQNGYVNNC-HVNNCMFTDQNNNEYTCSENRSVHPGNG 353
DB 3512 TCGPHE---FOCKN-----NNCIPDHWR-----DSQN-----DCSDNSDEENCKPQT 3551
QY 354 CSMQDRYVSGFCEVRERCKWCHMDKXNGTD 385
DB 3552 CTLKDFLCANGDC-VSSRFWCGDFDCADGSD 3582

RESULT 15

US-09-825-144-14
; Sequence 14, Application US/09825144
; Patent No. US20020037286A1
; GENERAL INFORMATION:
; APPLICANT: Matthias Krause
; APPLICANT: Antonio S. Sechi
; APPLICANT: Frank B. Gertler
; APPLICANT: Jorgen Weiland
; TITLE OF INVENTION: Methods for Altering T-Cell and Macrophage Activation
; FILE REFERENCE: M0656/7065
; CURRENT APPLICATION NUMBER: US/09/825,144
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/194,215
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 819
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-825-144-14

Query Match 3.2%; Score 103; DB 10; Length 819;

Best Local Similarity 19.5%; Pred. No. 5.6;
Matches 144; Conservative 76; Mismatches 236; Indels 284; Gaps 33;

QY 15 EVTSEPRCKKLKSTESVYFHNHNSNADFHRIOEKTG-----NDWVPVTII 60
DB 72 EKPEKEPPPLKPTGGSPRGCTQNSVRDPEVKVGLKPVSPKPTSLTKEDSKPVVLR 131
QY 61 DV--RGHSYLOENKTKTDLRLPHDEMGNRPDVIES-----IDSQVLQEA-- 105
DB 132 PPGNKLHNLNQESDLKTP-----GPKPGPAPPVPENELKPGFSKVAGAKSKFMPAAQD 184
QY 106 -----RPPLVSADDEIYSTK-----AFIGPIYKPPPEKKRN 137
DB 185 TDSKPRFPRTFTGQKPSLSTEDSOEENTSKNPVQKSPVOLGAKSKGAPFKPKEDPED 244
QY 138 EGRNEAH-----VLNGINDRG-----QKEKOKFENSEKSEIDNELFQ-----FYK 177
DB 245 KDHGAPSSPFGVVLKPAASRGSGGLSKNFEEKEDKTDLAKNIFLNKLNQEPAREPPK 304
QY 178 ETEL-----EKEKDFENCKESESQEFVPEYEGHNNGLL-----KPDE----- 219
DB 305 APSLCTAGTPWGOSEKEGDKNS---ATPKQKALPPL-----SVLGPFPKPNRPNNVD 355
QY 220 ----EKKDLSNKAMP--SHCDYQ-QNLGNEDKYPNGQVPTFCDTSTSFREWQSVYPF 274

DB 356 LTRFRKADSANSATKSQTPYSTTSLPPPPPTHPAS-----QPPLPA 396
QY 275 IVPYGPPLPSLNYHLNIQRFSGPPNPSPNIFOAQDSQIQNGYVNNCHVNNCMFTDQ 334
DB 397 SUPAHPVPFSL-----PPRNIKPPDLKHPIN-----DEN 426
QY 335 NEYTDSENRSVHPSGNGCSMDRYVSG--FCEVRERCKWCHMDKXNGTDNDRFVNOQF 392
DB 427 QD-----GYMHSQGTG-NLEEEQESGETYEDIDSSKERDKKREKEKKLELERKE 477
QY 393 QBEKLNKLOKLLILLRLGLPG-----SGKTTLSRILLQNRD-GIVFSTD- 435
DB 478 OKERERKEQELKKFK-LTGPIQVIHAKACCDVKGKNELS---FKQGEDIEIIRITDN 533
QY 436 -----DYFHHQDGYRYNNVNLG-----DAHDWNNQRAK 463
DB 534 PEGKWLGRTAGSYGIKTTAVEIDI-----DSLKRKKNLSNAVPPRLVEDDDQDYYDDVAE 589
QY 464 QAI-----DQGRSPVIIDNTNIOAWENKPYVEVA 492
DB 590 QDAPNSHGSGSGGMFPPTDDEIYDGEIEEDDDGSGYQVDEKTNANSWG- ---LKWL 646
QY 493 ICKGYRVEFHEPETWKKFDPBEELEKRNKHGVSRRKTAQMLDRYEQMSISIVNNSVEPSH 552
DB 647 KGKDDRKKSIRES-----PKVSESDNNEGSLPSQHKQLDVGEE-----VYDDVDASD 694
QY 553 KSTQRPQPPQGRQWGGSLG 572
DB 695 F-----PPPPAEMSQMSVG 709

Search completed: March 10, 2003, 18:04:49
Job time : 27 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 17:48:59 ; Search time 20 seconds
(without alignments)
857.678 Million cell updates/sec

Title: US-09-924-654-4

Perfect score: 3209
Sequence: 1 MSYGEIEGKFLGPREVTSE.....RQRMGSGSLGSHNRVCVTNNH 583

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3209	100.0	583	2	US-08-616-392C-4
2	449.5	14.0	132	2	US-08-616-392C-10
3	372	11.6	104	2	US-08-616-392C-12
4	127	4.0	394	4	US-09-195-188-1
5	117	3.6	769	3	US-08-725-459B-38
6	112	3.5	769	3	US-08-725-459B-40
7	110	3.4	769	3	US-08-725-459B-39
8	109.5	3.4	1147	3	US-08-470-260-5
9	109.5	3.4	1147	3	US-08-471-491-5
10	109.5	3.4	1147	4	US-08-466-662-5
11	109.5	3.4	1865	1	US-08-588-985-2
12	109.5	3.4	1865	1	US-08-971-988-2
13	109.5	3.4	3289	2	US-08-477-451-2
14	109	3.4	769	3	US-08-725-459B-37
15	108.5	3.4	2391	2	US-08-446-855A-2
16	108.5	3.4	2391	2	US-09-150-741-2
17	108	3.4	769	3	US-08-725-459B-35
18	107.5	3.3	859	1	US-08-053-614-2
19	107.5	3.3	859	1	US-08-316-397B-2
20	107.5	3.3	859	1	US-09-034-306-2
21	107.5	3.3	859	4	US-09-259-437-2
22	107.5	3.3	859	5	PCT-US93-09782-2
23	107.5	3.3	1181	1	US-08-053-614-4
24	107.5	3.3	1181	1	US-08-316-397B-4
25	107.5	3.3	1181	2	US-09-034-306-4
26	107.5	3.3	1181	4	US-09-259-437-4
27	107.5	3.3	1181	5	PCT-US93-09782-4

28	107.5	3.3	1939	4	US-09-310-187A-1	Sequence 1, Appl
29	107	3.3	630	3	US-08-725-459B-17	Sequence 17, Appl
30	107	3.3	680	3	US-08-725-459B-3	Sequence 3, Appl
31	107	3.3	730	3	US-08-725-459B-2	Sequence 2, Appl
32	107	3.3	733	3	US-08-725-459B-29	Sequence 29, Appl
33	107	3.3	733	3	US-08-725-459B-30	Sequence 30, Appl
34	107	3.3	773	1	US-08-524-757-6	Sequence 6, Appl
35	107	3.3	773	3	US-08-725-459B-1	Sequence 1, Appl
36	107	3.3	773	3	US-08-725-459B-79	Sequence 79, Appl
37	107	3.3	1507	3	US-08-929-329-5	Sequence 5, Appl
38	106.5	3.3	10182	4	US-09-134-001C-3159	Sequence 3159, Ap
39	106	3.3	506	2	US-08-820-170A-19	Sequence 19, Appl
40	106	3.3	506	3	US-09-055-699-19	Sequence 19, Appl
41	106	3.3	506	4	US-09-273-565-19	Sequence 19, Appl
42	106	3.3	506	4	US-09-565-538-19	Sequence 19, Appl
43	106	3.3	506	4	US-09-661-468-19	Sequence 19, Appl
44	105	3.3	324	2	US-08-816-755-2	Sequence 2, Appl
45	105	3.3	324	4	US-09-090-673-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-616-392C-4
Sequence 4, Application US/08616392C
Patent No. 5998165
GENERAL INFORMATION:
APPLICANT: Goold, Richard D.
APPLICANT: Akerblom, Ingrid E.
APPLICANT: Selhamer, Jeffrey
APPLICANT: Coleman, Roger
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES PANCIA
TITLE OF INVENTION: AND PANC1B ASSOCIATED WITH PANCREATIC CANCER
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/616,392C
FILING DATE: 15-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/581,240
FILING DATE: 29-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0052-1US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 583 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: hnt
CLONE: 486071
US-08-616-392C-4
Query Match 100.0%; Score 3209; DB 2; Length 583;
Best Local Similarity 100.0%; Pred. No. 1e-284;

```
Matches 583; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSYGEIEGFLGPREVTSPPRCKKLLKSTTSYVFHNSNADFHRIOEKTGNDWVPVTII 60
DB 1 MSYGEIEGFLGPREVTSPPRCKKLLKSTTSYVFHNSNADFHRIOEKTGNDWVPVTII 60
QY 61 DVRGHSYLOENKIKTTDLRLHDEMPGNRPDVIESIDSOVLQEARPPPLVSADDEIYSTS 120
DB 61 DVRGHSYLOENKIKTTDLRLHDEMPGNRPDVIESIDSOVLQEARPPPLVSADDEIYSTS 120
QY 121 KAFIGIYKPPKKRNEGRNEAHLVINGIDRGGOKEKQFNSEKSEIDNELFQFYKETE 180
DB 121 KAFIGIYKPPKKRNEGRNEAHLVINGIDRGGOKEKQFNSEKSEIDNELFQFYKETE 180
QY 181 ELEKEKDFENSCKESEPQOFVFPYEGHNGNLLKPDDEKDLNKNKAMPSCDYOQNLG 240
DB 181 ELEKEKDFENSCKESEPQOFVFPYEGHNGNLLKPDDEKDLNKNKAMPSCDYOQNLG 240
QY 241 NEPDKYPCNGQVPTFCDTSTSPRPMQSVYFIVPGPPLPSLNYHLNIORESGPPNP 300
DB 241 NEPDKYPCNGQVPTFCDTSTSPRPMQSVYFIVPGPPLPSLNYHLNIORESGPPNP 300
QY 301 PSNIFOAQDSDIQNGYYVNNCHVNNMNCMTFDQNNETDCSENRSVHPSGNCSMDQRY 360
DB 301 PSNIFOAQDSDIQNGYYVNNCHVNNMNCMTFDQNNETDCSENRSVHPSGNCSMDQRY 360
QY 361 VSGFCEVRERCKWCHMDKXNGTDRFVNOQFQEBKLNKQLKLLILLRGLPGSGKTTLSR 420
DB 361 VSGFCEVRERCKWCHMDKXNGTDRFVNOQFQEBKLNKQLKLLILLRGLPGSGKTTLSR 420
QY 421 ILLGNRQDGIIVSTDFYFHQDGYRYNVLGDADHWNONRAKQAIQDGRSPVIIDNTNI 480
DB 421 ILLGNRQDGIIVSTDFYFHQDGYRYNVLGDADHWNONRAKQAIQDGRSPVIIDNTNI 480
QY 481 QAWEMKPYVEVAIGRYRVEFHEPETWKKFDPPELEKRNKHGVSRRKKAQMLDRYEYQMS 540
DB 481 QAWEMKPYVEVAIGRYRVEFHEPETWKKFDPPELEKRNKHGVSRRKKAQMLDRYEYQMS 540
QY 541 ISIVNSVEPSHKSTQRPDPPOGRMGSGSLGSHNRVCVTNNH 583
DB 541 ISIVNSVEPSHKSTQRPDPPOGRMGSGSLGSHNRVCVTNNH 583

RESULT 2
US-08-616-392C-10
; Sequence 10, Application US/08616392C
; Patent No. 5998165
; GENERAL INFORMATION:
; APPLICANT: Goold, Richard D.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Seilhamer, Jeffrey
; APPLICANT: Coleman, Roger
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES PANCL1A
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/616,392C
; FILING DATE: 15-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/581,240
; FILING DATE: 29-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0052-1US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 104 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0052-1US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 132 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 533948
US-08-616-392C-10
```

```
Query Match 14.0%; Score 449.5; DB 2; Length 132;
Best Local Similarity 82.1%; Pred. No. 1.7e-33;
Matches 87; Conservative 0; Mismatches 0; Indels 19; Gaps 1;
QY 1 MSYGEIEGFLGPREVTSPPRCKKLLKSTTSYVFHNSNADFHRIOEKTGNDWVPVTII 60
DB 46 MSYGEIEGFLGPRE-----HHNSNADFHRIOEKTGNDWVPVTII 86
QY 61 DVRGHSYLOENKIKTTDLRLHDEMPGNRPDVIESIDSOVLQEAR 106
DB 87 DVRGHSYLOENKIKTTDLRLHDEMPGNRPDVIESIDSOVLQEAR 132
```

```
RESULT 3
US-08-616-392C-12
; Sequence 12, Application US/08616392C
; Patent No. 5998165
; GENERAL INFORMATION:
; APPLICANT: Goold, Richard D.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Seilhamer, Jeffrey
; APPLICANT: Coleman, Roger
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES PANCL1A
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/616,392C
; FILING DATE: 15-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/581,240
; FILING DATE: 29-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0052-1US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 104 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

US-08-616-392C-12

US-08-616-392C-12

Query Match	11.6%	Score 372;	DB 2;	Length 104;
Best Local Similarity	94.6%	Pred. No. 1.4e-26;		
Matches 70; Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0

Qy	Db
1	31
MSVGEIGKRLGREGVETSEPRCKRLKSTTESYVFNHNSNDFRRLQEKGNMVPVTII	MSVGEIGKRLGREGVETSEPRCKRLKSTTESYVFNHNSNDFRRLQEKGNMVPVTII
60	90
Qy	Db
61	91
DVRGHSLYLENFKTR 74	DVRGHSLYLENFKTR 104
IIIIIIIIIIII	IIIIIIIIIIII

RESULT 4
ITS-09-19

Sequence 1, Application US/09195188

```

: Patent NO. 6365/122
:
: GENERAL INFORMATION:
:
: APPLICANT: Marcu, Kenneth B.
: TITLE OF INVENTION: Y2H14 AN IKK BINDING PROTEIN
: FILE REFERENCE: 178-264
: CURRENT APPLICATION NUMBER: US/09/195,188
: CURRENT FILING DATE: 1998-11-17
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: PatentIn Ver. 2.0
:
: SEQ ID NO 1
:
: LENGTH: 394
:
: TYPE: PRN
:
: ORGANISM: Homo sapiens
:
: US-09-195-188-1

```

Query Match 4.0%; Score 127; DB 4; Length 394;
Best Local Similarity 25.5%; Pred. No. 0.0028;
Matches 41; Conservative 32; Mismatches 58; Indels 30; Gaps 7

```

0Y 401 OKLILGLGSGKTTSLRLLQNRD-----GIVSTDYF-----HHQD-----442
Db 81 ERIIVIMGLGSGKTHAKKILRIKREVEFEGGAPRVSLSDYFTTEYEKEEKDDSGKV 140
0Y 443 -----GYRYNQLDADHDMQNRNA-KOALIOGRSP-VIIDNTNQAMKPYEVAIGK 495
Db 141 KKKVDEYEVA-EMETVRTSMFKTFKTLDDGFFPFLITDAINDRVYVHNPQVSAATK 195
0Y 496 GYRVFHEBPETWKKPDEPELEKRNKHGVSRKIAQMLDRYE 536
Db 200 GEEVYLAEKSA---DNQTCGRNTHGKKLEINKMADHWE 236

```

RESULT 5
US-08-725-459B-38

Sequence 38, Application 05/08725459B
Patent No. 6084068
GENERAL INFORMATION:
APPLICANT: CONAWAY, RONALD C.
APPLICANT: CONAWAY, JOAN W.
TITLE OF INVENTION: ELONGIN A AND C FUNCTIONAL DOMAINS
NUMBER OF SEQUENCES: 79
CORRESPONDENCE ADDRESS:
ADDRESSEE: STIDLER & AUSTIN
STREET: 717 N. HARWOOD, SUITE 3400
CITY: DALLAS
STATE: TX
COUNTRY: US
ZIP: 75201-6507
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/725,459B
FILING DATE: 04-00E-1006

FILED DATE: 04-01-1990
CLASSIFICATION: 530

CLASIFICACION: 250

NAME: HANSEN, EUGENIA S.

REGISTRATION NUMBER: 31,

REFERENCE/DOCKET NUMBER: 11146/07501

TELECOMMUNICATION INFORMATION:

TELEPHONE: 214-981-3300
TELEFAX: 214-981-3400

TELEFAX: 214-961-3400

SEQUENCE CHARACTERISTICS:

LENGTH: 769 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: protein

NAME/REV. Protocol

NAME/KEY: P10CELL
LOCATION: 1 769

OTHER INFORMATION: /note

OTHER INFORMATION: 561-7

725-459B-38

Match 3.68

Local Similarity 19.3%; Pred. No. 0.063;
Local Mismatches 48; Indels 122; Gaps 16;
Local Conserved 79; Conservative 160;

```
QY      9 KFLGREGVSEPRCKLKSTTEZYVHNHNSMDFNRIQKGT- NMDVPITIIDVRGSX 67
        | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     307 KLTSPALDVASDNHFKPKHKDEKKIKSDKNKOSVDSDGRGTGPDLPAKKDKVPPNNLK 366
QY      68 LOEKIKITDTHREPLHEMGNGRPDIYEISIDSVLOEARPELVSADDELITYSKAFIGPI 127
        | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     367 AOEKVYTNS-----DKRSGSLPKVBE-----DMDDDEFQPTMSFESYL 407
QY      128 -YRPEKKRRNEGNENAHVLNGINDRGCGKEKOKFNSEKSEIDNELFOFKELIELEKEK 166
        | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     408 SYDPRRKKKKVVKTSGTAL-----GEKGLKKRKDKSTS-----KNLSAK-- 449
QY      187 DGFENSEKESPQEQEVPFYEGHNNGILKPDEEKKOLSNKAMPSHCDYOQLNEDPKY 246
        | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     450 -----LPKANEXSD-- KLDPAGAE-----PTRRKV 474
QY      247 PCNGOVIPTECDTSFTSRPQMOSVYPFIYVGPRLPSLYNLHNIQ---RFGSPMPSPS 302
        | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     475 PTD-VLPALPDIDLPALIQTNXR-----PLPSLETISSFORKRAKFSSPOEEE 521
QY      303 NIFDAQDDSOIQINGIYVNNCHVNANNCMTFPDONNEYTDCSEKRSSHVGSG----- 351
        | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     522 AGFGRGRNRNSKMQVYSGSKAYLRKMMTLHO-----OCIRVIDSIFFVGVPYSLEPVL 576
QY      352 NGCSTMOPY-----VSNGFCEVERECGDCHDXKDANGTDRFVNOOFOEEK 396
        | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     577 ECRTPDOLIREECNHVLIETDLMKYVC-----HRPFKEER 614
```

RESULT 6
US-08-725-459B-40

```

: Sequence 40, Application US/08/25459B
: Patent No. 6084068
:
: GENERAL INFORMATION:
:
: APPLICANT: CONAMAY, RONALD C.
: APPLICANT: CONAMAY, JOAN W.
: TITLE OF INVENTION: ELONGIN A AND C FUNCTIONAL DOMAINS
:
: NUMBER OF SEQUENCES: 79
:
: CORRESPONDENCE ADDRESS:
:
: ADDRESSEE: SIDLEY & AUSTIN
: STREET: 717 N. HARWOOD, SUITE 3400
: CITY: DALLAS
: STATE: TX
: COUNTRY: US
: ZIP: 75201-6507

```


GENERAL INFORMATION:

APPLICANT: Covacci, Antonello
 APPLICANT: Bugnoli, Massimo
 APPLICANT: Telford, John
 APPLICANT: Macchia, Giovanni
 APPLICANT: Rappuoli, Rino
 TITLE OF INVENTION: Helicobacter Pylori Proteins Useful
 TITLE OF INVENTION: for Vaccines and Diagnostics
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Chiron Corporation
 STREET: 4560 Horton Street
 CITY: Emeryville
 STATE: California
 COUNTRY: USA
 ZIP: 94608-2916
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/470,260
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/256,848
 FILING DATE: 21-OCT-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: McClung, Barbara G.
 REGISTRATION NUMBER: 33,113
 REFERENCE/DOCKET NUMBER: 0316, 001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (510) 601-2708
 TELEFAX: (510) 655-3542
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1147 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-470-260-5

Query Match 3.4%; Score 109.5; DB 3; Length 1147;
 Best Local Similarity 19.1%; Pred. No. 0.57;
 Matches 128; Conservative 98; Mismatches 240; Indels 205; Gaps 33;

QY 3 YGEIGKFLGPREVYSEPRCKKLKSTESYVFHNHNSADHRIOEKTGNDMPVPTIIDV 62
 DB 473 YGKKADKALDREKNVTLQGSILKH-----DGVAFVDYSNFKYTN-ASKNPNGVGT----- 522
 QY 63 RGHSTYLOE-NKIR-----TTDLRPLHDEMPGRNPVIESIDSOVLQEARP 108
 DB 523 NGVSHLEVGFKVAIFNLPLDNLNLAITSFVRNLEDKL-----TTKLSLQSEANKKLKD 576
 QY 109 LVSADDEIYSTSKAF-----IGPIYKPEPKKRNNGRNEAHVLAINGINDRG 153
 DB 577 FLSSNKELVGKTLNFKNAVADAKNTGNYDEVKAKQDKLEKSLRKREHLEKEVEKKLESKS 636
 QY 154 GQKEKQ-----KFNSEKSEI-----DNELFOFYKEIELEKE-KDGFENSKSESEP 198
 DB 637 GKNKMEAKAQAQNSQKDEIFALINKKANRDAIRAIAVQNLGKIKRELSDKLENVKNKLKD 696
 QY 199 SOEOFPPEYEGHNNGLLPDEEKDLSNKAMPSHCDYOQNLGNPDYKPCNGVYIPTFCD 258
 DB 697 FDKSFDEKRNKNDKFAETLAKLGSV-----KDLG----- 730
 QY 259 TSFTSFREMOGYVPFIYVPGPPLPSLVNHLNIOFSGPPNP-SNIFOADDDSOIQNGY 317
 DB 721 -----INDEWIS-----KVENLNAALN--EFKNGKNDKFSKYQAKSD--LENS- 770
 QY 318 YVNNCHVMMNCTFPDQNNETDSCENRBSVHSGNGSCMDQRYVSNCGCEVNERCWMKDH 377

DB 771 -VKDVII-----NQKVTDKVDNLNQ-----AVSVAKATGFSRVEQA-----L 807
 QY 378 MDKHNGTDREVNQOFE-EKLNKLQKLLI-----LRG-LPGSKTTLIRILLGQNRD 429
 DB 808 ADLKFNFSKEQLAQQAQKESLIMARKSEIYQSVKNGVGTGLVGLSQAETTLISKNSF 867
 QY 430 IVFSTDDYFHHODGYRRNV-OLGDADHWNQNAKQADIDGRSPVILIDNTNIOA-----W 483
 DB 868 I-----KELNAKLGSENNNNNGLK-----NEPIYAKVKKAKGAQASL 907
 QY 484 EMKPYEVA-----IGK--GYRVEFHEPETWKFDEPELEKRNKG 522
 DB 908 EEPYAGVAKKVNKIDRLNQLASGLVGQAAGFLPKRH-----DKYDLSKYG 957
 QY 523 VSR-KKIAQMLDREYQMS-----ISTVANSVEPSKSTQRP---PGQRQMGSLG 572
 DB 958 LSRNQLAOKIDNLNQAVSEAKQFQNLBOTIDKLDSKHNPMNLVESAKKVPASLS 1017
 QY 573 SHNVCVTNNH 583
 DB 1018 AKLDNYATNSH 1028

RESULT 9

US-08-471-491-5
 Sequence 5, Application US/08471491B
 Patent No. 6090611

GENERAL INFORMATION:

APPLICANT: Covacci, Antonello
 APPLICANT: Bugnoli, Massimo
 APPLICANT: Telford, John
 APPLICANT: Macchia, Giovanni
 APPLICANT: Rappuoli, Rino
 TITLE OF INVENTION: Helicobacter Pylori Proteins Useful For Vaccines And
 TITLE OF INVENTION: Diagnostics
 FILE REFERENCE: CHIR0044
 CURRENT APPLICATION NUMBER: US/08/471,491B
 CURRENT FILING DATE: 1995-06-06
 NUMBER OF SEQ ID NOS: 8
 SOFTWARE: Patentin Ver. 2.1
 SEQ ID NO 5
 LENGTH: 1147
 TYPE: PRT
 ORGANISM: Helicobacter pylori
 US-08-471-491-5

Query Match 3.4%; Score 109.5; DB 3; Length 1147;
 Best Local Similarity 19.1%; Pred. No. 0.57;
 Matches 128; Conservative 98; Mismatches 240; Indels 205; Gaps 33;

QY 3 YGEIGKFLGPREVYSEPRCKKLKSTESYVFHNHNSADHRIOEKTGNDMPVPTIIDV 62
 DB 473 YGKKADKALDREKNVTLQGSILKH-----DGVAFVDYSNFKYTN-ASKNPNGVGT----- 522
 QY 63 RGHSTYLOE-NKIR-----TTDLRPLHDEMPGRNPVIESIDSOVLQEARP 108
 DB 523 NGVSHLEVGFKVAIFNLPLDNLNLAITSFVRNLEDKL-----TTGLSLQSEANKKLKD 576
 QY 109 LVSADDEIYSTSKAF-----IGPIYKPEPKKRNNGRNEAHVLAINGINDRG 153
 DB 577 FLSSNKELVGKTLNFKNAVADAKNTGNYDEVKAKQDKLEKSLRKREHLEKEVEKKLESKS 636
 QY 154 GQKEKQ-----KFNSEKSEI-----DNELFOFYKEIELEKE-KDGFENSKSESEP 198
 DB 637 GKNKMEAKAQAQNSQKDEIFALINKKANRDAIRAIAVQNLGKIKRELSDKLENVKNKLKD 696
 QY 199 SOEOFPPEYEGHNNGLLPDEEKDLSNKAMPSHCDYOQNLGNPDYKPCNGVYIPTFCD 258
 DB 697 FDKSFDEKRNKNDKFAETLAKLGSV-----KDLG----- 730
 QY 259 TSFTSFREMOGYVPFIYVPGPPLPSLVNHLNIOFSGPPNP-SNIFOADDDSOIQNGY 317

Db 731 -----INPEWIS-----KVENLNAALN--EEKNGKNGKDFSKVTOAKSD--LENS- 770
QY 318 YVNNCHVNNMCTFDONNEYTDCSENRSVHPGSGCNSMODRVYVNGFCVEVRERCKWCHC 377
Db 771 -VKDVII-----NOKVTDKVDNLQ-----AVSVAKATGDFSRVEQA-----L 807
QY 378 MDKHNGTDRFVNOQFOE-EKLNKLQKLLIL-----LRG-LPGSGKTTLSRILLGONRDG 429
Db 808 ADLNKFSKQLAQAQAKNESLNARKKSEIYQSVKNGVNGTLYVNGLSQAQAEATLSKNFSD 867
QY 430 IVESTDDYFHODGYRYNVN-QLGDAHDWNQNAKQAIQDGRSPVITDNTNQA-----W 483
Db 868 I-----KKELNAKLGNFNNNNGLK-----NEPIYAKVNNKKKAGQAASL 907
QY 484 EMKPYVEVA-----IGK--GYRVEFHEPETWKFDPPELEKRNKKG 522
Db 908 EEPYIAQVAKVNAKIDRLNQIASGLGVVQAGAPFLKRH-----DKVDDLKSYG 957
QY 523 VSR-KKIAOMLDRYEQMS-----ISIVNNSVEPSHKSTQRPDP--PQGRQWCGSLG 572
Db 958 LSRNQELAQKIDNLQAVSEAKAGFGNLEQITDKLKDSTKHPNMLWVESAKKVPASLS 1017
QY 573 SHNRVCVTNNH 583
Db 1018 AKLDNYATNSH 1028
RESULT 10
US-08-466-662-5
; Sequence 5, Application US/08466662B
; Patent No. 6130059
; GENERAL INFORMATION:
; APPLICANT: Covacci, Antonello
; APPLICANT: Bugnoli, Massimo
; APPLICANT: Telford, John
; APPLICANT: Macchia, Giovanni
; APPLICANT: Rappuoli, Rino
; TITLE OF INVENTION: Helicobacter Pylori Proteins Useful For Vaccines And
; FILE REFERENCE: CHIR0057
; CURRENT APPLICATION NUMBER: US/08/466,662B
; CURRENT FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1147
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-08-466-662-5
Query Match 3.4%; Score 109.5; DB 4; Length 1147;
Best Local Similarity 19.1%; Pred. No. 0.57;
Matches 128; Conservative 98; Mismatches 240; Indels 205; Gaps 33;
QY 3 YGEIEGKFLGPREEVTSERPKLSTTESYVPHNNSADFHRIQKTDNDWVPVITIDV 62
Db 473 YGKADKALDREKNVTLOGSLKH-----DGVMFVDYSNFKYTN-ASKNPKGVGT----- 522
QY 63 RGHSLQOE--NKIK-----TTDLHRPLHDEMFGNRPDVTESIDSVLQEARPP 108
Db 523 NGVSHLEVGFNKVAIFNLPDLNLAITSFVRNLEDKL-----TTKGUSPOEANKLIND 576
QY 109 LVSADDEIYSTKAF-----IGPIYKPEKKRNEGRNEAHVLNGINDRG 153
Db 577 FLSSNKELVGTILNFNAKAVADAKNTGNYDEVKKAQKDELSLRKREHLEKEVEKKLESKS 636
QY 154 GQKEQ-----KNSEKSEI-----DNELFQYKEIELEKE-KDGFNSCKSESP 198
Db 637 GKNKNEAKAQAQNSQKDEIFALINKEANDRAIAIAQNLKGIKRELSKLENVNKNLKD 696
QY 199 SOBQFVPFYEGHNGLLKDPDEEKDLSNKAHPSHCDYQOQNLGNPDYKPCNGOVPTTFC 258
Db 697 FDKSDFEFGKNGKDFSKAETILKALGVS-----KDLG----- 730

QY 259 TSFTSRPRQSWQYPIVPGPPLPSLNHLNQRSGPPNPP-SNIFOAQDDSQIQNGY 317
Db 731 -----INPEWIS-----KVENLNAALN--EFKNGKNGKDFSKVTOAKSD--LENS- 770
QY 318 YVNNCHVNNMCTFDONNEYTDCSENRSVHPGSGCNSMODRVYVNGFCVEVRERCKWCHC 377
Db 771 -VKDVII-----NOKVTDKVDNLQ-----AVSVAKATGDFSRVEQA-----L 807
QY 378 MDKHNGTDRFVNOQFOE-EKLNKLQKLLIL-----LRG-LPGSGKTTLSRILLGONRDG 429
Db 808 ADLNKFSKQLAQAQAKNESLNARKKSEIYQSVKNGVNGTLYVNGLSQAQAEATLSKNFSD 867
QY 430 IVESTDDYFHODGYRYNVN-QLGDAHDWNQNAKQAIQDGRSPVITDNTNQA-----W 483
Db 868 I-----KKELNAKLGNFNNNNGLK-----NEPIYAKVNNKKKAGQAASL 907
QY 484 EMKPYVEVA-----IGK--GYRVEFHEPETWKFDPPELEKRNKKG 522
Db 908 EEPYIAQVAKVNAKIDRLNQIASGLGVVQAGAPFLKRH-----DKVDDLKSYG 957
QY 523 VSR-KKIAOMLDRYEQMS-----ISIVNNSVEPSHKSTQRPDP--PQGRQWCGSLG 572
Db 958 LSRNQELAQKIDNLQAVSEAKAGFGNLEQITDKLKDSTKHPNMLWVESAKKVPASLS 1017
QY 573 SHNRVCVTNNH 583
Db 1018 AKLDNYATNSH 1028
RESULT 11
US-08-588-985-2
; Sequence 2, Application US/08588985
; Patent No. 5777094
; GENERAL INFORMATION:
; APPLICANT: Michiyuki MATSUDA et al.
; TITLE OF INVENTION: cDNA OF DOCK180 GENE AND DOCK180 PROTEIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,985
; FILING DATE: January 19, 1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1865 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL:

ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE: spleen cell of homo sapiens
US-08-588-985-2

Query Match 3.4%; Score 109.5; DB 1; Length 1865;
Best Local Similarity 18.7%; Pred. No. 1.2;
Matches 139; Conservative 97; Mismatches 277; Indels 229; Gaps 35;

```

QY 4 GEIEGKFL---GPREVTSEPRCKKLTSTESVFNH-----HSNADF-----43
D 1074 GQHKIKFPEWVGPILLETILPETELRKATIP--IFEDMOCERHSTSFQMFENEITK 1131
QY 44 --HRIQKTGDMVPVITIDV-----RGHSYLO---ENKIKTT-----DLHRPLDE 85
D 1132 LDHEVEGGRGEOYKVLFDKILLEHCRKHXYLAKGTGFVKLVYLMRLLDYRTIMDE 1191
QY 86 MGRNPDIYESIDSOVLOEARRPPLVSADDEIYSTSKAFIGPYKPEKKRNEGRNE---142
D 1192 NKENRMCSTVAVLNPKYEIER-----EMV-----IRLYKCLDLHKEDNYTEAY 1238
QY 143 ---AHYLANGIND-----RGQOKEK-----OKFNSEK-----165
D 1239 TLLHAKLKSEDEVCAVHNLQROGYQATTGQLKEQLYQELIHFDGKKMEEAIALGK 1298
QY 166 ---SEIDNELFOFYKEIELEKENDGFENSCSESPSOEOPVPEYEGH-----NGLL--215
D 1299 ELAQYENEMFDYEQLSLELKKQAQFYEINIVKIRKPDYFVAGYGGGFPFLGKVF 1358
QY 216 ---KPEDEKDLKSNKAMSHCDYQO-----NLGNEPDKPCNGQVITPFCSTSTS 263
D 1359 YRGKEYPEDEFEARLLQFPNAEKMTTSPGGDIKNSP-----GOYIOCF-----1406
QY 264 FRPEWQSVYPIVYGPPLPSLNYHLN-IORF-----SGPPNPSPNIFQAOODSOJON 315
D 1407 VKPKLDLPKPHRPVSEQIVSF-YRVNEVQREYISRPTRKGEKNPDNEFAMMIRITTYT 1465
QY 316 GYVYNNCHVNN-----NCM-TFQDNEEYTDSCSENNSSVHPS--GNGCSM-356
D 1466 TAYKLPGLRMEYKSVMEISPLENAIETMQLNDKINSVWQOHLDDPSLPRLSML 1525
QY 357 ---QDRVYNGCFEVRRCRCKDKCHMDKNGTDPRVNOQFOEEKLNKQOKLL-----ILL 407
D 1526 LNGVIDPVMGFAFYEAFTDRLQENPRA-----HEKTEKLDLAWOIPFLA 1576
QY 408 RGLPQSG-KTTLRILLQONRDGIYFSTDYFNHODGYRYVNNOLDGANDMNONAKOAI 466
D 1577 EGIRHGDKVTEDALRPFHERMEACQOLKEKYEKYGVRIMPSSLDDBRG---SRPRGMV 1633
QY 467 DQGRSP-----VIIDNTNIQAMEMKRYVEVALGKGYRVEFHEPETW---KEDPE 513
D 1634 RSFTPPSSSRPLSVASVSLSDSTSPSRGSD---GFALEPLRLPKMHSKQDKLKD 1688
QY 514 ELEKRNKIGVGRKKLAQMLDREYGM---STSI-----VNN-546
D 1689 DLEKAKKKEKKEKRSKQOIEFEKEKFTDLSQOSEAVIILSETISPLRPRKQVAMNV 1748
QY 547 -----SVPSHKSTQRRPPP 561
D 1749 IGSERRFVSPPSSPSQOTPPP 1770

```

RESULT 12
US-08-971-988-2
Sequence 2, Application US/08971988
Patent No. 5786461
GENERAL INFORMATION:
APPLICANT: MICHIOYUKI MATSUDA et al.
TITLE OF INVENTION: CDNA OF DOCK180 GENE AND DOCK180 PROTEIN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Menderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700

CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/971,988
FILING DATE: 17-NOV-1997
CLASSIFICATION: 536

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/588,985
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:

TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 1865 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL:

ANTI-SENSE:
FRAGMENT TYPE:

ORIGINAL SOURCE: spleen cell of homo sapiens
US-08-971-988-2

Query Match 3.4%; Score 109.5; DB 1; Length 1865;
Best Local Similarity 18.7%; Pred. No. 1.2;
Matches 139; Conservative 97; Mismatches 277; Indels 229; Gaps 35;

```

QY 4 GEIEGKFL---GPREVTSEPRCKKLTSTESVFNH-----HSNADF-----43
D 1074 GQHKIKFPEWVGPILLETILPETELRKATIP--IFEDMOCERHSTSFQMFENEITK 1131
QY 44 --HRIQKTGDMVPVITIDV-----RGHSYLO---ENKIKTT-----DLHRPLDE 85
D 1132 LDHEVEGGRGEOYKVLFDKILLEHCRKHXYLAKGTGFVKLVYLMRLLDYRTIMDE 1191
QY 86 MGRNPDIYESIDSOVLOEARRPPLVSADDEIYSTSKAFIGPYKPEKKRNEGRNE---142
D 1192 NKENRMCSTVAVLNPKYEIER-----EMV-----IRLYKCLDLHKEDNYTEAY 1238
QY 143 ---AHYLANGIND-----RGQOKEK-----OKFNSEK-----165
D 1239 TLLHAKLKSEDEVCAVHNLQROGYQATTGQLKEQLYQELIHFDGKKMEEAIALGK 1298
QY 166 ---SEIDNELFOFYKEIELEKENDGFENSCSESPSOEOPVPEYEGH-----NGLL--215
D 1299 ELAQYENEMFDYEQLSLELKKQAQFYEINIVKIRKPDYFVAGYGGGFPFLGKVF 1358
QY 216 ---KPEDEKDLKSNKAMSHCDYQO-----NLGNEPDKPCNGQVITPFCSTSTS 263
D 1359 YRGKEYPEDEFEARLLQFPNAEKMTTSPGGDIKNSP-----GOYIOCF-----1406
QY 264 FRPEWQSVYPIVYGPPLPSLNYHLN-IORF-----SGPPNPSPNIFQAOODSOJON 315
D 1407 VKPKLDLPKPHRPVSEQIVSF-YRVNEVQREYISRPTRKGEKNPDNEFAMMIRITTYT 1465
QY 316 GYVYNNCHVNN-----NCM-TFQDNEEYTDSCSENNSSVHPS--GNGCSM-356
D 1466 TAYKLPGLRMEYKSVMEISPLENAIETMQLNDKINSVWQOHLDDPSLPRLSML 1525

```

Qy 357 ---QDRYVSNCFCEVRCHKDHCMKNGTDRFVNOQFOEKLKLOKLL-----ILL 407
Db 1526 LNVIVDPAMVGGFANYERKAFFTDRLQHPHA-----HEKIEKLDLIAWQIPFLA 1576
Qy 408 RGLPGSG-KTTLRSLILQNGRDGIVFSTDDYFHHQDGYRYNVNOLGDADHWNQRAKQAI 466
Db 1577 EGIRHKGVTALRPPHFERMEACFKQLKEVEYGVRIIMPSSLDLDRG---SRPRSMV 1633
Qy 467 DQGRSP-----VIIDNTNIQWEMKPYVEVAIGKGYRVEFHEPFW-----KEDPE 513
Db 1634 RSFTMPSSRPLSVASVSSLDSTPSRPGSD-----GFALEPLIPKKMHSRQDKLKD 1688
Qy 514 ELEKRNHGVSRKKAQMLDREYQM-----SISI-----VNN- 546
Db 1699 DLEKXDKKKEKRNKSHQEIFEKEFKPTDISLOOSEAVILSETISPLRPORPKSQVMNV 1748
Qy 547 -----SVEPSHKSQTORPPPP 561
Db 1749 IGSERREFSVSPSSSQOTPPP 1770
RESULT 13
US-08-477-451-2
; Sequence 2, Application US/08477451
; Patent No. 5928865
; GENERAL INFORMATION:
; APPLICANT: Covacc1, Antonello
; TITLE OF INVENTION: Helicobacter Pylori Cagi Region
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,451
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0335.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2708
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3289 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-477-451-2
Query Match 3.4%; Score 109.5; DB 2; Length 3289;
Best Local Similarity 19.1%; Pred. No. 2.9;
Matches 128; Conservative 98; Mismatches 240; Indels 205; Gaps 33;
Qy 3 YGEIEGFLGPREVTSEPRCKKLKSTESVFHNHNSADFHRIQETGNDWVPTIIDV 62
Db 2004 YGKADRALDREKNVTLQGLKH-----DGMVFVDYNSFKYTN-ASKNPNKGVGT---- 2053
Qy 63 RGHSLQOE--NKIK-----TTDLRPLHDEMGNRPDVFIESIDSVLQEARPP 108
Db 2054 NGVSHLEVGNKVAIFNLPLDNLNLAITSFVRNLEDKL-----TTKGLSPQEAANKLIK 2107

Qy 109 LVSADDEIYSTSKAF-----LGPYKPPKPKKRNKGRNEAHVLNGINDRG 153
Db 2108 FLSSNKELVGTTLFNLKAVADAKNTGNYDEVKKAQKDLKSLRKRREHLEKEVEKLESKS 2167
Qy 154 GOKKQKQ-----KFNEKSEI-----DNELFOFYKEIEBELEKE-KDGFENSKSESEP 198
Db 2168 GNKKMEAKAQAQKQDEIFALINKENRDRARATAYAGNLKGIKRELSDKLENNVKNLKD 2227
Qy 199 SQEQVFYFEGHNGLLKPDEEKDLSNKAHPSHCDYQONLGNPDYKPCNGQVPTFC 258
Db 2228 FDKSDFEFGKNGKDFSKAEETLKALGVS-----KDLG----- 2261
Qy 259 TSFTSFRPWSQSVFPIVPGPPLSLNYHNIORFSGPPNPP-SNIFQAQDDSOIQNGY 317
Db 2262 -----INPEWIS-----KVENLNAALN--EFKNGKNKDFSKVTOAKSD--LENS- 2301
Qy 318 YVNNCHVNNWNTFDQNNEXYDCSENSSVHPVSGNGCSMODRYVSYNGFCEVRERCKWDHC 377
Db 2302 -VKDVII-----NOKVTDKVDNLNQ-----AVSAKATGDFSRVEQA-----L 2338
Qy 378 MDKHGTDREYVNOQFOE-EKLNKLOKLLIL-----LRG-LPGSGKTTLSRILLQONRDG 429
Db 2339 ADLANFSKEQLAQQAQAKNESLNARKKSEIYQSVKNGVNGTIVGNGLSQAEATTLKKNFSD 2398
Qy 430 IVFSTDDYFHHQDGYRYNVN-OLGDAHDWNNQRAKQAIQDGRSPVIIDNTNIA-----W 483
Db 2399 I-----KKELNAKLGNNNNNGLK-----NEPIYAKVNNKKAGQAASL 2438
Qy 484 EMKPYVEVA-----IGK--CYRVEFHEPETWKFDPPELEKRNKHG 522
Db 2439 EEPYIAQVAKVKNNAKIDRLNOLASGLGVVGOAGFPLKRH-----DKVDDLSKVG 2488
Qy 523 VSR-KKIAQMLDREYQMS-----ISTVMNSVPSHKSQTORPPP-----PQGRQWGGSLG 572
Db 2489 LSRNQLAQKIDNLNOAVSEAKAGFFGNLEQITDKLKDSTKHNPNNLWVESAKKVPASLS 2548
Qy 573 SHNRVCVTNNH 583
Db 2549 AKLDNYATNSH 2559
RESULT 14
US-08-725-459B-37
; Sequence 37, Application US/08725459B
; Patent No. 6084068
; GENERAL INFORMATION:
; APPLICANT: CONAWAY, RONALD C.
; APPLICANT: CONAWAY, JOAN W.
; TITLE OF INVENTION: ELONGIN A AND C FUNCTIONAL DOMAINS
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SIDLEY & AUSTIN
; STREET: 717 N. HARWOOD, SUITE 3400
; CITY: DALLAS
; STATE: TX
; COUNTRY: US
; ZIP: 75201-6507
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/725,459B
; FILING DATE: 04-OCT-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: HANSEN, EUGENIA S.
; REGISTRATION NUMBER: 31,966
; REFERENCE/DOCKET NUMBER: 11146/07501
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 214-981-3300
; TELEFAX: 214-981-3400

INFORMATION FOR SRO ID NO: 37:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 769 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: not relevant
 MOLECULE TYPE: protein
 FEATURE:
 NAME/KEY: Protein
 LOCATION: 1..769
 OTHER INFORMATION: /note="amino acids 1-552 and
 OTHER INFORMATION: 557-773 of Elongin A"
 US-08-725-459B-37

Query Match 3.4%; Score 109; DB 3; Length 769;
 Best Local Similarity 19.6%; Pred. No. 0.34;
 Matches 79; Conservative 48; Mismatches 165; Indels 112; Gaps 16;

Db 9 KFLGPREVTSSEPRCKLKSTTESYVFHNSNADFHRIQKGTG-NDWVPTIIDRGHSY 67
 307 KKLSPALDVASDNHFKFKHKOSEKIKSDKKOSVDSYSGRGTDPLPRAKDKVPPNLIK 366
 QY 68 LOENKIKITDLHRLHDEMPGNRPDIYESIDSOYLQEARPLVSADDEIYSTSKAFICPI 127
 Db 367 AQDEKVRINS-----DKRSPGLPKVEE-----MDMDDEFEGPTMSFESYL 407
 QY 128 -YRPEKKRNEGRNEAHVLNGINDRGQKQKQFNSEKSEIDELFOFYIEIELEKEK 186
 Db 408 SYDOPRKKKKVVTSGTAL-----GEKGLKRRKSKTS-----KNLSAOK-- 449
 QY 187 DGFENSCSESPSOQFPFPEFEGHNGILKPEDEKDKLSNKAAPSHDQOQNLGNEPKY 246
 Db 450 -----LPKANENKSD--KLQPGAE-----PRPRKY 474
 QY 247 PCNGVITFCDSFTSEPRMOSYPRIVYGPPLPSLNYHLNQ-----RFGPPNPS 302
 Db 475 PTD--VLPLALDIPLAQTNYR-----PLPLELISSTQPRKAFSSQDEEE 521
 QY 303 NIFQAQDSQIQNGIYVNNCHVNNCMTFDQ--NNEYIDCE--NRSSVHPSGNGCSM 356
 Db 522 AGFGRANSMKQVYSGSGCAYLPKPMTLHQVKKNNIDISIFEVGVPYSVLEPVERCTP 581
 QY 357 QDRY-----VSNGFCEVRERCKDHGMKHNDRFVNOQPOEEK 396
 Db 582 DQLYRIECCNHVLIETDQMKVHC-----HRDFKEER 614

RESULT 15
 US-08-446-855A-2
 Sequence 2, Application US/08446855A
 Patent No. 5849573
 GENERAL INFORMATION:
 APPLICANT: Stewart, Thomas S
 APPLICANT: Flores, Maria V
 TITLE OF INVENTION: Nucleotide sequence encoding carbamoyl
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESSES:
 ADDRESS: Nixon & Vanderhye PC
 STREET: 1100 No. 5849573th Glade Road, 8th Floor
 CITY: Arlington
 STATE: Virginia
 COUNTRY: USA
 ZIP: 22201-4714
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.24
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/446,855A
 FILING DATE: 06-Jul-1995

CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Mitchard, Leonard C
 REGISTRATION NUMBER: 29,009
 REFERENCE/DOCKET NUMBER: 47-80
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-816-4000
 TELEFAX: 703-816-4100
 INFORMATION FOR SRO ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2391 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-446-855A-2

Query Match 3.4%; Score 108.5; DB 2; Length 2391;
 Best Local Similarity 17.2%; Pred. No. 2.2;
 Matches 105; Conservative 80; Mismatches 190; Indels 235; Gaps 24;

QY 2 SYGIEGKFLGPREVTSSEPRCKLKSTTESYVFHNSNADFHRIQKGTGNDWVPTIID 61
 Db 1500 TYKKNKYKDMGLNNNIYDELNGTSHSTNDHLYLDNFTSD-----EIGNN----- 1546
 QY 62 VRGHSYLOENKIKITDLH-----RPLHDEMPGNRPDIYES-----IDSOYLQ 103
 Db 1547 -----KNNDMLSEKSTISNKNPGNSIYVDSYVNNKIKMKMELIDNEULN 1594
 QY 104 EAPPLVSADDEIYSTSKAFICPIYKPEKKRNEGRNEAHVLNGINDRGQKQKQFN 163
 Db 1595 DEYNNNVNMNSNYNNASAFV-----NGKDRND-----NL 1624
 QY 164 EKSEIDELFOFYIEIELEKEKDFENSC-----KESPSOQFPFPEFEGH-NGGLLPD 218
 Db 1625 ENDCIEKNMDHTYKHYNRLNNRSTNERMMLVNNKESNHE-----KGHRNGLANKN 1678
 QY 219 EEK-----KOLSNKAMPSH-CDYOQNLGNEPKDKPCNGOVIPTFCDSFTSFRPEW 268
 Db 1679 KEKNNEKKNKKNKDKKNHYNNHKNNEYSNNISKEF--NNYV--DDINKKYEIDE 1732
 QY 269 QSVYPTI--VPYGPPLPSLNY--HLNIQFIS-----GPPNPSNIFQAQ 308
 Db 1733 NDIYFTSSQGNNDLNDNTLSSELTNDIYDDDDYYDEDEDDYDDDDDDDDDD 1792
 QY 309 DDSQIQNGIYVNNCHVNNCMTFDQNNEXYDQSENRSVHPSGNGCSMODRYSNGFCEV 368
 Db 1793 GEDEEDNDYNDGDYSNLSRSSRIDY-----SSVIYSGNENIFNEKINDIGFKII 1845
 QY 369 REBCWKD-----HCHDKHNGTDRFVNOQPOEEKLKL 400
 Db 1846 DNRNEKEKEKKCFYLGCGCYRIGSSVEFDSAHCY-----TIRKL 1889
 QY 401 OKLLILLGLPGSGKTTLSR-----ILL--GONRGIVF 432
 Db 1890 NKKAILINCNPETVSTVDYDESRLLYFDEITTEVIFITNFENSNGVITIAFGQTSNNLVF 1949
 QY 433 STDYFHHQDGYRYVNOGLGAH-----DNQ--NRAQA 465
 Db 1950 SL-----YKNNVILGSVHKVLIYKIGINFRTYVILKIDQPKNNKFKLSKAIOF 2000
 QY 466 IDQGRSPVIT 475
 Db 2001 ANEVKFPVLY 2010

Search completed: March 10, 2003, 17:55:29
 Job time : 28 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 17:39:28 ; Search time 42 seconds
(without alignments)
1849.646 Million cell updates/sec

Title: US-09-924-654-4

Perfect score: 3209
Sequence: 1 MSYGEIEGKFLGPREVTSE.....RQRMGSLSGSHNVCYTNNH 583

Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	3209	100.0	583	18 AAN26654 Human PANCJA assoc
2	3205	99.9	594	21 AAB56501 Human prostate can
3	542	16.9	119	23 ABB42992 Human ovarian anti
4	509	15.9	659	22 AAB93057 Human protein sequ
5	425.5	13.3	243	22 AAB84326 Amino acid sequenc
6	334	10.4	969	22 ABB63009 Drosophila melanog
7	321.5	10.0	127	22 AAU28356 Novel human secret
8	247	7.7	137	22 AAU28168 Novel human secret
9	195.5	6.1	227	19 AAN98859 H. pylori GHPO 166
10	139.5	4.3	726	22 AAG6716 Human transcriptio

11	130.5	4.1	2485	21 AAB18172 Plasmodium falci
12	128	4.0	2500	17 AAB18272 Plasmodium falci
13	128	4.0	2913	17 AAM00384 Plasmodium falci
14	127	4.0	394	23 ABB76271 Human IKK binding
15	127	4.0	629	22 AAU72168 Human RNA metaboli
16	122	3.8	799	21 AAB21003 Human nucleic acid
17	121.5	3.8	1123	21 AAU58277 Heliothis armigera
18	121.5	3.8	2010	21 AAB18218 Plasmodium falci
19	120	3.7	1080	22 ABB11277 Novel human diagno
20	120	3.7	1844	21 AAU37362 Staphylococcus aur
21	117.5	3.6	702	23 ABB04333 Plasmodium falci
22	117	3.6	759	19 AAB60796 Human zinc finger
23	117	3.6	942	22 ABB62498 Deletion mutant de
24	117	3.6	1383	22 ABB63586 Drosophila melanog
25	116	3.6	797	22 ABB62042 Drosophila melanog
26	115.5	3.6	1247	21 AAB18215 Plasmodium falci
27	115.5	3.6	1714	21 AAB18275 Plasmodium falci
28	115	3.5	1156	22 ABB26549 Novel human diagno
29	113.5	3.5	1817	21 AAB18301 Plasmodium falci
30	113.5	3.5	417	22 AAG70780 S cerevisiae apopt
31	113	3.5	507	21 AAB18173 Plasmodium falci
32	113	3.5	980	21 AAB18294 Plasmodium falci
33	113	3.5	64	22 ABB16674 Novel human diagno
34	112.5	3.5	1979	21 AAB18171 Plasmodium falci
35	112.5	3.5	769	19 AAB60798 Deletion mutant de
36	112	3.5	1064	22 ABB59430 Drosophila melanog
37	111.5	3.5	662	22 AAM41117 Human polyptide
38	111.5	3.5	2295	21 AAB18320 Plasmodium falci
39	111.5	3.5	904	21 AAB18370 Plasmodium falci
40	111	3.5	1166	21 AAB18268 Plasmodium falci
41	111	3.5	1308	21 AAB18167 Plasmodium falci
42	110.5	3.4	939	22 ABB58143 Drosophila melanog
43	110.5	3.4	1373	22 ABB62368 Drosophila melanog
44	110.5	3.4	552	22 ABB59326 Drosophila melanog
45	110			

ALIGNMENTS

RESULT 1	AAW26654	AAW26654 standard; Protein; 583 AA.
ID	AAW26654	
XX	AAW26654	
AC	16-FEB-1998	(first entry)
XX		
DE	Human PANCJA associated with pancreatic cancer.	
XX		
KW	PANCJA; pancreas cancer; diagnosis; therapy; human.	
XX		
OS	Homo sapiens.	
XX		
PN	WO9724437-A1.	
XX		
PD	10-JUL-1997.	
XX		
PF	18-DEC-1996;	96W0-US19981.
XX		
PR	15-MAR-1996;	96US-0616392.
XX		
PR	29-DEC-1995;	95US-0581240.
XX		
PA	(INCY-) INCYTE PHARM INC.	
XX		
PI	Aberblom IE, Coleman R, Goold RD, Seilhamer JJ;	
XX		
DR	WPI: 1997-363673/33.	
XX		
DR	N-PSDB; AAT91010.	
XX		
PT	DNA encoding PANCJA and PANCIB associated with pancreatic cancer -	
XX	useful for diagnosis of pancreatic cancer and identifying risk of	
PT	cancer	

XX PS Disclosure; Page 46-47; 65pp; English.

XX CC This protein comprises human PANC1A, which is associated with

CC pancreatic cancer and possibly other cancers. Its sequence was

CC deduced from a full-length PANC1A cDNA clone (see AAT91010) derived

CC from human teratocarcinoma hNT-2 cells. Specific fragments of

CC PANC1A (see AAT91008) and PANC1B (see AAT91009) cDNA are claimed.

CC Also claimed are: antisense molecules; expression vectors;

CC transformed host cells; and purified PANC1A and PANC1B

CC polypeptides. Diagnostic testing for the presence of PANC1A or

CC PANC1B in a sample provides a means for positively correlating

CC abnormal levels of PANC1A or PANC1B with pancreatic cancer

CC (claimed). The host cells can be cultured for the recombinant

CC production of PANC1A or PANC1B proteins (claimed). PANC1A and

CC PANC1B polypeptides can be used to screen for antagonists and

CC inhibitors, and to raise diagnostic antibodies.

XX CC

XX SQ Sequence 583 AA;

Query Match 100.0%; Score 3209; DB 18; Length 583;

Best Local Similarity 100.0%; Pred. No. 8,9e-255;

Matches 583; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSYGEIEGKFLGPREEVTSERPRCKKLKSTTSYVFHNHNSADFHRIQKTDNDWVPVTII 60

Db 1 MSYGEIEGKFLGPREEVTSERPRCKKLKSTTSYVFHNHNSADFHRIQKTDNDWVPVTII 60

Qy 61 DVRGHSYLOENKIKTTDLRPLHDEMPCNRPDVIESDSQVLQEARPLVSADDEIYSTS 120

Db 61 DVRGHSYLOENKIKTTDLRPLHDEMPCNRPDVIESDSQVLQEARPLVSADDEIYSTS 120

Qy 121 KAFIGPIYKPEKKRNEGRNEAHLVNGINDRGQKQKFNSEKSEIDNELFQFYKEIE 180

Db 121 KAFIGPIYKPEKKRNEGRNEAHLVNGINDRGQKQKFNSEKSEIDNELFQFYKEIE 180

Qy 181 ELEKEKDGFCNSKESPSQSQFVPFYEGHNGLLKPDDEKKDLSNKAHPSHCDYQOQNLG 240

Db 181 ELEKEKDGFCNSKESPSQSQFVPFYEGHNGLLKPDDEKKDLSNKAHPSHCDYQOQNLG 240

Qy 241 NEPKYPCNGQVITPFCDTSTSPRPQSQSVYPIVPGPLPSLNYHLNTORESGPPNP 300

Db 241 NEPKYPCNGQVITPFCDTSTSPRPQSQSVYPIVPGPLPSLNYHLNTORESGPPNP 300

Qy 301 PSNIFQADDSQIQNGYYVNNCHVNNWNCMTFDQNNYTDGSENSSVHPSGNGCSMDRY 360

Db 301 PSNIFQADDSQIQNGYYVNNCHVNNWNCMTFDQNNYTDGSENSSVHPSGNGCSMDRY 360

Qy 361 VSNGFCEVRERCWKDCHCKKNGTDRFVYNQOFOEKLNLKLLGLLGLPGSGKTTLSR 420

Db 361 VSNGFCEVRERCWKDCHCKKNGTDRFVYNQOFOEKLNLKLLGLLGLPGSGKTTLSR 420

Qy 421 ILLGQNRDGIYFSTDYFHHODGYRYNNVNLGDAHDNQNRAKQIDQGRSPVIIDNTNI 480

Db 421 ILLGQNRDGIYFSTDYFHHODGYRYNNVNLGDAHDNQNRAKQIDQGRSPVIIDNTNI 480

Qy 481 QAWEMKPYVEVAIGKRYVEFHEPETWKKFDPDELEKRNKKGVRKKAQMLDRYEYQMS 540

Db 481 QAWEMKPYVEVAIGKRYVEFHEPETWKKFDPDELEKRNKKGVRKKAQMLDRYEYQMS 540

Qy 541 ISIVMNSVEPSHKSTORPPPPQGRQWGGSLGSHNRVCVTNNH 583

Db 541 ISIVMNSVEPSHKSTORPPPPQGRQWGGSLGSHNRVCVTNNH 583

RESULT 2

AAB56501

ID AAB56501 standard; Protein: 594 AA.

XX AC AAB56501;

XX AC

DT 13-MAR-2001 (first entry)

XX XX

DE Human prostate cancer antigen protein sequence SEQ ID NO:1079.

XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;

KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;

KW vulnary; gastrointestinal; nephrotropic; antineoplastic; gynaecological;

KW antibacterial; gene therapy; neural; immune; reproductive; renal;

KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;

KW wound; infectious disease.

XX Homo sapiens.

OS WO2000055174-A1.

PN 21-SEP-2000.

PD 08-MAR-2000; 2000WO-US05988.

PF 12-MAR-1999; 99US-0124270.

PR (HUMA-) HUMAN GENOME SCI INC.

PA (ROSE/) ROSEN C A.

XX Rosen CA, Ruben SM;

XX WPI: 2000-587513/55.

DR N-PSDB: AAF15704.

PT Prostate cancer associated gene sequences, referred to as prostate

PT cancer antigens, useful for treatment, prevention, and diagnosis of

PT disorders such as prostate cancer -

XX Claim 11: Page 1502-1504; 2338pp; English.

XX AAF15566 to AAF16505 encode the human prostate cancer associated

CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.

CC The prostate cancer antigens can have neuroprotective, cytostatic,

CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,

CC nephrotropic, antineoplastic, gynaecological and antibacterial activities,

CC and can be used in gene therapy. The prostate cancer antigen

CC polynucleotides may be used for detection of prostate cancer, chromosome

CC identification, as chromosome markers, and for numerous other diagnostic

CC or research purposes. The prostate cancer antigens may be used to treat

CC disorders such as neural, immune, muscular, reproductive,

CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative

CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to

CC AAB57303 represent sequences used in the exemplification of the present

CC invention.

XX SQ Sequence 594 AA;

Query Match 99.9%; Score 3205; DB 21; Length 594;

Best Local Similarity 99.8%; Pred. No. 1,9e-254;

Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSYGEIEGKFLGPREEVTSERPRCKKLKSTTSYVFHNHNSADFHRIQKTDNDWVPVTII 60

Db 12 MSYGEIEGKFLGPREEVTSERPRCKKLKSTTSYVFHNHNSADFHRIQKTDNDWVPVTII 71

Qy 61 DVRGHSYLOENKIKTTDLRPLHDEMPCNRPDVIESDSQVLQEARPLVSADDEIYSTS 120

Db 72 DVRGHSYLOENKIKTTDLRPLHDEMPCNRPDVIESDSQVLQEARPLVSADDEIYSTS 131

Qy 121 KAFIGPIYKPEKKRNEGRNEAHLVNGINDRGQKQKFNSEKSEIDNELFQFYKEIE 180

Db 132 KAFIGPIYKPEKKRNEGRNEAHLVNGINDRGQKQKFNSEKSEIDNELFQFYKEIE 191

Qy 181 ELEKEKDGFCNSKESPSQSQFVPFYEGHNGLLKPDDEKKDLSNKAHPSHCDYQOQNLG 240

Db 192 ELEKEKDGFCNSKESPSQSQFVPFYEGHNGLLKPDDEKKDLSNKAHPSHCDYQOQNLG 251

Qy 241 NEPKYPCNGQVITPFCDTSTSPRPQSQSVYPIVPGPLPSLNYHLNTORESGPPNP 300

Db 252 NEPKYPCNGQVITPFCDTSTSPRPQSQSVYPIVPGPLPSLNYHLNTORESGPPNP 311

XX Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis and treatment of
 PT cancer, neurological, inflammatory, and autoimmune disorders -
 XX
 PS Example 2: SEQ ID No 713; 107pp; English.
 XX
 CC The invention relates to novel isolated human secreted polypeptides (I)
 CC and polynucleotides (II). (I) and (II) are useful for treating
 CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
 CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
 CC involved in increasing haematopoiesis, stem cell survival, bone growth
 CC and remodeling. (I), (II) and modulators of (II) are useful for
 CC prophylaxis or treatment of one or more cancers. (II) is also useful for
 CC creating transgenic animals useful for studying the in vivo activities of
 CC the polypeptide as well as for studying modulators of the polypeptides.
 CC (I) induces the proliferation of neural cells and regeneration of nerve
 CC and brain tissue and is useful for the treatment of central and
 CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
 CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
 CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth, and in tissue repair, healing of burns, incisions,
 CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
 CC disorders, or periodontal disease. Furthermore, (I) is also useful for
 CC gut protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues, various immune deficiencies and
 CC disorders including severe combined immunodeficiency (SCID), bacterial or
 CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
 CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
 CC reactions and conditions, such as asthma or other respiratory problems.
 CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
 CC fertility, metabolism, catabolism, anabolism, storage or elimination of
 CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
 CC analgesic effects or other pain reducing effects, immunoglobulin like
 CC activity and can act as an antigen in a vaccine composition to raise an
 CC immune response. AAU28020-AAU28395 represent novel human secreted protein
 CC amino acid sequences of the invention.
 XX
 SQ Sequence 127 AA;

Query Match 10.0%; Score 321.5; DB 22; Length 127;
 Best Local Similarity 53.9%; Pred. No. 1.6e-18;
 Matches 62; Conservative 18; Mismatches 34; Indels 1; Gaps 1;

Qy 398 NKLOKLLILLRGLPGSGKTTLSRILLGNRDGIVFSTDYFHODG-YRYNNVNLGDAHD 456

Db 12 HSFRKHLVLLRGLPGSGKTTLSRILLGNRDGIVFSTDYFHODG-YRYNNVNLGDAHD 71

Qy 457 WNQRAKQADQGRSPVITDNTNQAWMKPYVEVAIGKGYRVEFPEPTWKFED 511

Db 72 WNQKRXAXIARNGISPIIIDNTNLHAWMKPYAVMALENNYEVIFREDPTRWKEN 156

RESULT 8

AAU28168

ID AAU28168 standard; Protein; 137 AA.

AC AAU28168;

XX AAU28168;

DT 18-DEC-2001 (first entry)

XX Novel human secretory protein. Seq ID No 337.

XX Human, secreted protein; arthritis; Crohn's disease; sepsis; shock;
 KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;

KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 KW fertility; analgesic; pain; antigen.
 XX Homq sapiens.
 OS WO200166689-A2.
 PN 13-SEP-2001.
 XX 05-MAR-2001; 2001WO-US04942.
 XX 07-MAR-2000; 2000US-0519705.
 XX 19-MAY-2000; 2000US-0574454.
 PR 17-JUN-2000; 2000US-0596193.
 PR 14-JUL-2000; 2000US-0616847.
 PR 19-SEP-2000; 2000US-0665363.
 PR 20-OCT-2000; 2000US-0693267.
 PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
 PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
 XX WPI: 2001-589934/66.

DR N-PSDB: AAS45068.

XX Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis and treatment of
 PT cancer, neurological, inflammatory, and autoimmune disorders -
 XX

PS Example 5: SEQ ID No 337; 107pp; English.

XX The invention relates to novel isolated human secreted polypeptides (I)
 CC and polynucleotides (II). (I) and (II) are useful for treating
 CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
 CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
 CC involved in increasing haematopoiesis, stem cell survival, bone growth
 CC and remodeling. (I), (II) and modulators of (II) are useful for
 CC prophylaxis or treatment of one or more cancers. (II) is also useful for
 CC creating transgenic animals useful for studying the in vivo activities of
 CC the polypeptide as well as for studying modulators of the polypeptides.
 CC (I) induces the proliferation of neural cells and regeneration of nerve
 CC and brain tissue and is useful for the treatment of central and
 CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
 CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
 CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth, and in tissue repair, healing of burns, incisions,
 CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
 CC disorders, or periodontal disease. Furthermore, (I) is also useful for
 CC gut protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues, various immune deficiencies and
 CC disorders including severe combined immunodeficiency (SCID), bacterial or
 CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
 CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
 CC reactions and conditions, such as asthma or other respiratory problems.
 CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
 CC fertility, metabolism, catabolism, anabolism, storage or elimination of
 CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
 CC analgesic effects or other pain reducing effects, immunoglobulin like
 CC activity and can act as an antigen in a vaccine composition to raise an
 CC immune response. AAU28020-AAU28395 represent novel human secreted protein
 CC amino acid sequences of the invention.

XX Sequence 137 AA;

Query Match 7.7%; Score 247; DB 22; Length 137;
 Best Local Similarity 42.9%; Pred. No. 2.3e-12;
 Matches 48; Conservative 21; Mismatches 39; Indels 4; Gaps 1;

Qy 469 GRSPVITDNTNQAWMKPYVEVAIGKGYRVEFPEPTWKFEDPEELEKRNKHGVSRRKI 528

[illegible]

```

DB 1908 -----KGAGVYFSAMRSENVDFPNH--SFFENLN-----EKKKKYIDE----- 1946
OY 476 DNTNIQAMEMKRYEVAIGKGYVEFHEPFTWKKFDPPELF-----KRNKHGVSRRK 527
DB 1947 --TNN-----ENYESDNEYDSEDDDESDNDEQNKENERGDEKOG 1986
OY 528 IAQML-----DREYQMSISIVANSYSPSHS 554
DB 1987 YEEMNGDGKNEEMNGDKNE-EMNVGDKNGKGINEEHKN 2023

RESULT 12
AAB18272
ID AAB18272 standard; Protein: 2500 AA.
AC AAB18272;
XX
XX 07-NOV-2000 (first entry)
DE Plasmodium falciparum chromosome 2 related protein SEQ ID NO:129.
XX
XX Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
KW antimalarial; malaria; protozoacide; infection; insecticide.
XX
OS Plasmodium falciparum.
XX
XX WO200025728-A2.
XX
XX 11-MAY-2000.
PD
XX
XX 05-NOV-1999; 99WO-US26796.
PR
XX 05-NOV-1998; 98US-0107131.
PA (HOEFF/) HOFFMAN S.
PA (CARU/) CARUCCI D.
PA (GARD/) GARDNER M.
PA (VENT/) VENTER J C.
XX
XX Hoffman S, Carucci D, Gardner M, Venter JC;
XX
XX WPI: 2000-365347/31.
DR
XX
XX Proteins encoded by chromosome 2 of the human malarial parasite,
PT Plasmodium falciparum, useful as antimalarial vaccines and in the
PT diagnosis of P.falciparum infection.
XX
XX Disclosure: Page 302-309; 577pp; English.
PS
XX
XX The present invention describes proteins and their fragments (I) encoded
CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
CC Also described are: (I) nucleotide sequences (II) encoding (I); and (2)
CC vaccines against P. falciparum infection comprising (I) or (II).
CC (I) and (II) are useful for the development of vaccines against
CC P. falciparum infection. (I) and polyclonal antisera or a monoclonal
CC antibody raised to immunogens comprising the sequences of (I), are
CC useful in the detection of infection with P. falciparum. Furthermore,
CC (I) (especially when they are rifins or secreted or membrane proteins)
CC can aid the identification of drugs to treat or prevent P. falciparum
CC infection, or they can be used to identify drug resistance in
CC P. falciparum. Sequencing of the Plasmodium chromosome 2 and the
CC subsequent identification of proteins encoded by it will help to expand
CC our understanding of parasite biology, a process hampered by the
CC complexity of the parasitic lifecycle, and provide new targets for
CC vaccine and drug development. Parasite resistance to drugs and mosquito
CC resistance to insecticides have led to a resurgence of malaria in many
CC parts of the world, and there is a pressing need for vaccines and new
CC drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352 represent nucleotide
CC and protein sequences given in the present invention, but which are not
CC specifically mentioned within the specification.
XX
XX Sequence 2500 AA;

```

```

Query Match 4.0%; Score 128; DB 21; Length 2500;
Best Local Similarity 20.5%; Pred. No.1.1;
Matches 89; Conservative 68; Mismatches 187; Indels 90; Gaps 20.

OY 6 IEGKFLGPR---EETVSEPRCKKLKSTESYVFHNHNSADFRIOETKGNDAVPTIIVD 62
DB 1010 INGQYINDEYITSEETSE-----EYTSBGYINEGYINQYINGQYIN----- 1052
OY 63 RCHSYLQENKIK--TTDLHRLHDEMPGNRPDVIESIDQYLQEARPLYSADDEIYST 119
DB 1053 ---QYINGOSIEDOSTND--QSIDOSTNDQSIDOSTNDQSIDOS----TNDQSIINO 1103
OY 120 SKAFIGPIYKPEPKKRNENGRNEAHVLNINRGCGKEKQKNSKSEIENDLPPFYKEI 179
DB 1104 STNDQPPNEQPPNEQPPNESTEDQCLIDKNVNEOLS-----DQTLNRYF--- 1149
OY 180 ELEKEKDGFEKSKES-----EPSQEQFVPPEYEGHNGLLKPKDEKKDLKSKAMPS 231
DB 1150 --LECSQRTESLCNHSLSDPYMNNDNSHQYSNGETENDVLSQENPVVDDLSGHQNN 1207
OY 232 HCDYQONLQNEP-DKXPC---NGQYIPFPCD--TSFTSFRQWQSVYPFIVPGPPLPS 284
DB 1208 DNSFNSSSNVPLNVPNTIENSILPLIEGTNSAHLNFRGSYSDPFF--HSPNYSI 1264
OY 285 LNYHLNIOFSGPPNPSPSIFQADDSQIQNGYIYNKNCHVKNCM-----TFDON- 334
DB 1265 LEFCCS-RYFSS--NPFPEKTMIONE-QVQESLYISNFKANHYERIKITHIDPFTSN 1320
OY 335 ---NEYTD--CSENRSSVHPGNGCSMODRYVSGFCVRECRCKDHQMDHNG----T 384
DB 1321 LLYKNEITDKETISENKNKEIIEENKIIENKIVKKNENMYKKNVKNVKNENYK 1380
OY 385 DFEVNOQFOEERKLN 398
DB 1381 SRFIKKEHISMILN 1394

RESULT 13
AAM00384
ID AAM00384 standard; Protein: 2913 AA.
XX
XX AAM00384;
AC
XX
XX 21-FEB-1997 (first entry)
DE
XX
XX Plasmodium falciparum erythrocyte membrane protein.
XX
XX Plasmodium falciparum; erythrocyte membrane protein; malaria;
KW detection; identification; treatment; prevention; parasite.
XX
XX Plasmodium falciparum MC type.
OS
XX
XX Key Location/Qualifiers
FH Domain 62..394
FH Domain /label= Duffy binding ligand domain 1
FH Region 607..648
FH Domain /note= "Cysteine rich motif"
FH Domain 839..1272
FH Domain /label= Duffy binding ligand domain 2
FH Region 1482..1527
FH Domain /note= "Cysteine rich motif"
FH Domain 1706..2005
FH Domain /label= Duffy binding ligand 3
FH Domain 2102..2349
FH Region /label= Duffy binding ligand 4
FH Region 2354..2398
FH Domain /note= "Cysteine rich motif"
FH Domain 2450..2475
FH Domain /note= "Putative transmembrane domain"
XX
XX WO9633736-A1.

```

```
PD 31-OCT-1996.
XX
PF 26-APR-1996; 96WO-US05798.
XX
PR 27-APR-1995; 95US-0430908.
XX
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Baruch DI, Howard RJ, Pasloske BL;
XX
DR WPI; 1996-497376/49.
XX
DR N-PSDB; AAT41852.
XX
PT New Plasmodium falciparum erythrocyte membrane proteins - used to
PT develop products for the diagnosis, treatment or prevention of
PT malaria parasite infections
XX
PS Claim 1; Figure 12; 149pp; English.
XX
CC A polypeptide comprising a Plasmodium falciparum (Pf) erythrocyte
CC membrane protein 1 (PfEMP1) or active fragments or analogues of that
CC protein can be used in the treatment or prevention of symptoms of a
CC malaria parasite infection. The polypeptides can inhibit, block or
CC reverse the sequestration of erythrocytes in patients suffering from
CC malaria. Nucleic acids derived from the PfEMP1 gene can be used as
CC probes and primers to identify a Plasmodium falciparum parasite, the
CC primers used to generate characteristic amplification patterns from
CC different P. falciparum strains. Antibodies specifically
CC immunoreactive with the PfEMP1 polypeptide or its fragments may be
CC used in diagnosis of malaria infection. This is the PfEMP1 protein
CC of the MC type of Plasmodium falciparum. An alternative, truncated
CC PfEMP1 protein is given in AAW00385.
XX
SQ Sequence 2913 AA;
Query Match 4.0%; Score 128; DB 17; Length 2913;
Best Local Similarity 20.8%; Pred. No. 1.3; Indels 126; Gaps 22;
Matches 88; Conservative 56; Mismatches 153;
QY 117 YTSKAFIGPIYKPPKKRNGEARNHAHLNGIN----- 150
DB 1461 YNLLKHKIDPCI-----KREKQDKTEHKINGCNKCEVKRWLETKNGWGNKKHYNI 1515
QY 151 DRGQKQKQKFNSEKSEINELQF-YKEIEE-LEKEKGFENSEKSESPSEQFVPFYE 208
DB 1516 NSNDDKETIAYNKSYFVDQGLFDTDYKKAQKVVEDEK-----BRKKIWGCT 1562
QY 209 GHNNGLLKPDEKKDLNKAHPCHDYQNLGNEDPKYPCNGOVIPTFCDTSFSPPEW 268
DB 1563 GHDECEKEKENKNITNLI---SELQDKITSCQKNHPNGK---TACD-PFPSTPPE- 1614
QY 269 QSYVYPIVYGPPLPSNLNHLNIQFSGPPNPP-----SNIFQAQDSDSIQNGYVYNN 321
DB 1615 -ETDPLDDDDTDPDLD-DOHTQPKFCPPPPPPPTCVEXIAKELRVEAEGKINNELKNG 1672
QY 322 CHVWNCMTFDQNNETDCSENRSSVHPSGNGSMQDRVVSNGFCFVRCRWKDHCMKH 381
DB 1673 KDFNGKCNVKKNGKNAVIGES-----CKFEQTY-ENSVNNINNK-----KD 1714
QY 382 NGTDFE-VNQOFOEELNKLKLLI-----LRGLPGSGKTTLS----- 419
DB 1715 QNERFKIGQKNWFYIGIRKDLGIRPREHMCDDLSMLRTTISDALLKKTQEA 1774
QY 420 -----RILLQNRDGIIVSTDDYFHHQ--DGVRVYNNQLGD---AHD-WNQNRKOA 465
DB 1775 KSERDDIIRKLEQN-----SCDE---HRICDAMKYSADLGDIIIRGDLNKNKSKQG 1825
QY 466 IDQ 468
DB 1826 LQK 1828
RESULT 14
```

```
ABB76271
ID ABB76271 standard; Protein; 394 AA.
XX
AC ABB76271;
XX
DT 12-AUG-2002 (first entry)
XX
DE Human IKK binding protein Y2H14.
XX
KW Y2H14; IKK binding protein; I kappa B kinase; cytostatic;
KW antiinflammatory; immunosuppressive; immunostimulant; human.
XX
OS Homo sapiens.
XX
PN US365722-B1.
XX
PD 02-APR-2002.
XX
PF 17-NOV-1998; 98US-0195188.
XX
PR 17-NOV-1998; 98US-0195188.
XX
PA (UYNV ) UNIV NEW YORK STATE RES.
XX
PI Marcu KB;
XX
DR WPI; 2002-442213/47.
DR N-PSDB; ABL57342.
XX
PT Polynucleotide encoding an I kappa B kinase binding protein Y2H14 and
PT the recombinant protein encoded for elucidating and controlling
PT pathways leading to inflammation and apoptosis
XX
PS Claim 1; Column 11-14; 9pp; English.
XX
CC The present sequence is the protein sequence for a human I kappa B
CC kinase (IKK) binding protein, designated Y2H14. Y2H14 was shown,
CC using yeast two-hybrid screens, to bind to the C-terminal domain
CC regions of IKK-alpha and IKK-beta, specifically to the region of
CC the IKK proteins made up of the contiguous helix-loop-helix and
CC leucine zipper domains. The Y2H14 protein and other IKK binding
CC proteins are useful for elucidating and controlling pathways
CC leading to inflammation and apoptosis. They can also be used to
CC detect IKK complexes and modulate IKK activity in cells undergoing
CC signalling by inflammatory mediators, and to identify
CC therapeutically active agents that modulate the binding or
CC interaction of Y2H14 with IKK-alpha or -beta. Molecules that
CC prevent the formation of Y2H14/IKK complexes or inhibit the
CC dissociation of these complexes are useful for boosting the
CC immune system, or as immunosuppressives, or as antiinflammatory
CC agents.
XX
SQ Sequence 394 AA;
Query Match 4.0%; Score 127; DB 23; Length 394;
Best Local Similarity 25.5%; Pred. No. 0.082;
Matches 41; Conservative 32; Mismatches 58; Indels 30; Gaps 7;
QY 401 OKLLILLRGLPGSGKTTLSRILLGNRD-----GVFSTDDYF-----HHOD----- 442
DB 81 ERIVIMRGLPGSGKTHVAKLRDRKEVEFGGAPRVLSLDDYFITEVEKEEDPSGKKV 140
QY 443 -----GYRVNQLGDAHDWNNRRA-KQAIDQGRSP-VIIDNTNIQAWEMKPYVEAIGK 495
DB 141 KKKVREYEVA-EMETVYRTSMFKTFKTKLDDGFFPFIILDAINDRVHRFDQFWSAATK 199
QY 496 GYRVFEPHETWKKDPDELEKRNKRGVSRKKIAQMLDRIE 536
DB 200 GFEVYLAEMSA-----DNQTCGRNIHGRKLKEINKMADHWE 236
RESULT 15
AAI72168
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 18:04:15 ; Search time 10.733 Seconds
(without alignments)
55.005 Million cell updates/sec

Title: US-09-924-654-4_COPY_404_417
Perfect score: 69
Sequence: 1 LILLRGLPGSGKTT 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 188354 seqs, 42170167 residues

Total number of hits satisfying chosen parameters: 188354

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : Published_Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCT08_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	69	100.0	583	10 US-09-924-654-4	Sequence 4, Appl
2	69	100.0	594	10 US-09-925-300-1079	Sequence 1079, Ap
3	50	72.5	92	9 US-10-156-239-29	Sequence 29, Appl
4	50	72.5	92	9 US-10-156-239-30	Sequence 30, Appl
5	50	72.5	92	9 US-10-156-239-42	Sequence 42, Appl
6	50	72.5	92	10 US-09-795-693-29	Sequence 29, Appl
7	50	72.5	92	10 US-09-795-693-30	Sequence 30, Appl
8	50	72.5	92	10 US-09-795-693-42	Sequence 42, Appl
9	50	72.5	191	10 US-09-815-242-11348	Sequence 11348, A
10	50	72.5	235	10 US-09-803-286A-4	Sequence 64, Appl
11	50	72.5	479	10 US-09-971-309-64	Sequence 11562, A
12	49	71.0	448	10 US-09-815-242-11562	Sequence 5542, Ap
13	47	68.1	215	10 US-09-815-242-5542	Sequence 12227, A
14	47	68.1	224	10 US-09-815-242-12227	Sequence 12824, A
15	47	68.1	224	10 US-09-815-242-12824	Sequence 11948, A
16	47	68.1	457	10 US-09-815-242-11948	Sequence 10965, A
17	47	68.1	462	10 US-09-815-242-10965	Sequence 5755, Ap
18	47	68.1	547	9 US-09-738-626-5755	Sequence 6, Appl
19	46	66.7	65	10 US-09-756-996-6	

20	46	66.7	192	9 US-09-974-879-193	Sequence 193, App
21	46	66.7	453	10 US-09-815-242-10263	Sequence 10263, A
22	46	66.7	501	10 US-09-815-242-14089	Sequence 14089, A
23	46	66.7	748	10 US-09-949-192-43	Sequence 43, Appl
24	46	66.7	751	10 US-09-815-242-5832	Sequence 5832, Ap
25	46	66.7	825	10 US-09-815-242-12963	Sequence 12963, A
26	46	66.7	859	10 US-09-815-242-10623	Sequence 10623, A
27	46	66.7	891	10 US-09-949-192-45	Sequence 45, Appl
28	46	66.7	897	10 US-09-949-192-49	Sequence 49, Appl
29	45	65.2	212	10 US-09-815-242-13219	Sequence 13219, A
30	45	65.2	213	10 US-09-815-242-4889	Sequence 4889, Ap
31	45	65.2	231	10 US-09-815-242-10536	Sequence 10536, A
32	45	65.2	360	9 US-09-738-626-5332	Sequence 5332, Ap
33	45	65.2	459	9 US-09-738-626-5303	Sequence 5303, Ap
34	45	65.2	472	10 US-09-815-242-4945	Sequence 4945, Ap
35	45	65.2	477	10 US-09-815-242-10861	Sequence 10861, A
36	44	63.8	164	9 US-09-738-626-5292	Sequence 5292, Ap
37	44	63.8	191	9 US-09-895-913A-336	Sequence 336, App
38	44	63.8	213	10 US-09-815-242-10129	Sequence 10129, A
39	44	63.8	213	10 US-09-815-242-14070	Sequence 14070, A
40	44	63.8	215	10 US-09-815-242-5382	Sequence 5382, Ap
41	44	63.8	221	10 US-09-815-242-12195	Sequence 12195, A
42	44	63.8	307	9 US-10-108-605-183	Sequence 183, App
43	44	63.8	466	10 US-09-815-242-13525	Sequence 13525, A
44	44	63.8	523	10 US-09-815-242-13254	Sequence 13254, A
45	44	63.8	523	10 US-09-814-041A-2	Sequence 2, Appl

ALIGNMENTS

```
RESULT 1
US-09-924-654-4
; Sequence 4, Application US/09924654
; Patent No. US20020146712A1
; GENERAL INFORMATION:
; APPLICANT: Goold, Richard D.
; APPLICANT: Akedlorn, Ingrid E.
; APPLICANT: Selhamer, Jeffrey J.
; APPLICANT: Coleman, Roger
; TITLE OF INVENTION: TUMOR SUPPRESSOR
; FILE REFERENCE: PC-0049 CIP
; CURRENT APPLICATION NUMBER: US/09/924, 654
; CURRENT FILING DATE: 2001-08-07
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020146712A1 496071CD1
US-09-924-654-4

Query Match      100.0%; Score 69; DB 10; Length 583;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LILLRGLPGSGKTT 14
Db      404 LILLRGLPGSGKTT 417

RESULT 2
US-09-925-300-1079
; Sequence 1079, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
```

; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1079
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (430)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1079

Query Match 100.0%; Score 69; DB 10; Length 594;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LILLRGLPGSGKTT 14
|||||
Db 415 LILLRGLPGSGKTT 428

RESULT 3
US-10-156-239-29
; Sequence 29, Application US/10156239
; Publication No. US20030036074A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: No. US20030036074A1el Nucleic Acid Sequences Encoding Human Trans
; TITLE OF INVENTION: AtPase Molecule, A Human Ubiquitin Hydrolase-Like Molecule, A Hu
; TITLE OF INVENTION: Ubiquitin Conjugating Enzyme-Like Molecule, and Uses Therefor
; FILE REFERENCE: 35800/247645
; CURRENT APPLICATION NUMBER: US/10/156,239
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 09/795,693
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,906
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 09/809,557
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/192,018
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/808,568
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/191,790
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/808,767
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/191,781
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pfam consensus sequence
US-10-156-239-30

Query Match 72.5%; Score 50; DB 9; Length 92;
Best Local Similarity 71.4%; Pred. No. 0.65;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LILLRGLPGSGKTT 14
:||| |
Db 4 VLLVGPFGSGKTT 17

RESULT 5
US-10-156-239-42
; Sequence 42, Application US/10156239
; Publication No. US20030036074A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: No. US20030036074A1el Nucleic Acid Sequences Encoding Human Tr
; TITLE OF INVENTION: AtPase Molecule, A Human Ubiquitin Hydrolase-Like Molecule, A
; TITLE OF INVENTION: Ubiquitin Conjugating Enzyme-Like Molecule, and Uses Therefor
; FILE REFERENCE: 35800/247645
; CURRENT APPLICATION NUMBER: US/10/156,239
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 09/795,693
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,906
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 09/809,557
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/192,018
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/808,568
; PRIOR FILING DATE: 2001-03-14

Query Match 72.5%; Score 50; DB 9; Length 92;
Best Local Similarity 71.4%; Pred. No. 0.65;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LILLRGLPGSGKTT 14
:||| |
Db 4 VLLVGPFGSGKTT 17

```

; PRIOR APPLICATION NUMBER: 60/191,790
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/808,767
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/191,781
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 92
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pfam consensus sequence
US-10-156-239-42

Query Match
Best Local Similarity 71.4%; Score 50; DB 9; Length 92;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LILRLGPGSGKTT 14
Db 4 VVLLVGPFGSGKTT 17

RESULT 6
US-09-795-693-29
; Sequence 29, Application US/09795693
; Patent No. US20020068710A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 20685, 579, 17114, 23821, 33894, and
; FILE REFERENCE: 35800/209292
; CURRENT APPLICATION NUMBER: US/09/795,693
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,906
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 92
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pfam consensus sequence
US-09-795-693-29

Query Match
Best Local Similarity 72.5%; Score 50; DB 10; Length 92;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LILRLGPGSGKTT 14
Db 4 VVLLVGPFGSGKTT 17

RESULT 7
US-09-795-693-30
; Sequence 30, Application US/09795693
; Patent No. US20020068710A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 20685, 579, 17114, 23821, 33894, and
; FILE REFERENCE: 35800/209292
; CURRENT APPLICATION NUMBER: US/09/795,693
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,906
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
```

```

; LENGTH: 92
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pfam consensus sequence
US-09-795-693-30

Query Match
Best Local Similarity 72.5%; Score 50; DB 10; Length 92;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LILRLGPGSGKTT 14
Db 4 VVLLVGPFGSGKTT 17

RESULT 8
US-09-795-693-42
; Sequence 42, Application US/09795693
; Patent No. US20020068710A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 20685, 579, 17114, 23821, 33894, and
; FILE REFERENCE: 35800/209292
; CURRENT APPLICATION NUMBER: US/09/795,693
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,906
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 92
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pfam consensus sequence
US-09-795-693-42

Query Match
Best Local Similarity 72.5%; Score 50; DB 10; Length 92;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LILRLGPGSGKTT 14
Db 4 VVLLVGPFGSGKTT 17

RESULT 9
US-09-815-242-11348
; Sequence 11348, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
```

; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11348
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11348

Query Match 72.5%; Score 50; DB 10; Length 191;
Best Local Similarity 71.4%; Pred. No. 1.3;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LILLRGLPGSGKTT 14
| | | | | | | | | |
Db 4 LFLIAGPGSGKTT 17

RESULT 10
US-09-803-286A-4
; Sequence 4, Application US/09803286A
; Patent No. US20020062504A1
; GENERAL INFORMATION:
; APPLICANT: Tanksley, Steven D.
; APPLICANT: Brommonschenkel, Sergio H.
; TITLE OF INVENTION: PLANT GENE CONFERRING RESISTANCE TO TOSPOVIRUSES
; FILE REFERENCE: 19603/3201
; CURRENT APPLICATION NUMBER: US/09/803,286A
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 60/188,356
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Lycopersicon var.
US-09-803-286A-4

Query Match 72.5%; Score 50; DB 10; Length 235;
Best Local Similarity 64.3%; Pred. No. 1.7;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LILLRGLPGSGKTT 14
| | | | | | | | | |
Db 4 VISITGMPGSGKTT 17

RESULT 11
US-09-971-309-64
; Sequence 64, Application US/09971309
; Patent No. US20020106675A1
; GENERAL INFORMATION:
; APPLICANT: UEMORI, Takashi
; APPLICANT: SATO, Yoshiaki
; APPLICANT: FUJITA, Tomoko
; APPLICANT: MIYAKE, Kazuo
; APPLICANT: MURAI, Hiroyuki
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS
; FILE REFERENCE: 1422-0494P
; CURRENT APPLICATION NUMBER: US/09/971,309
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 09/446,504
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: PCT/JP98/02845

; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: JP 9-187496
; PRIOR FILING DATE: 1997-06-26
; PRIOR APPLICATION NUMBER: JP 9-320692
; PRIOR FILING DATE: 1997-11-27
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
US-09-971-309-64

Query Match 72.5%; Score 50; DB 10; Length 479;
Best Local Similarity 76.9%; Pred. No. 3.4;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ILLRGLPGSGKTT 14
: | | | | | | | | | |
Db 46 LLLAGPGSGKTT 58

RESULT 12
US-09-815-242-11562
; Sequence 11562, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.01A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11562
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11562

Query Match 71.0%; Score 49; DB 10; Length 448;
Best Local Similarity 64.3%; Pred. No. 4.5;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LILLRGLPGSGKTT 14
: | | | | | | | | | |
Db 96 VVLMAGLOGSGKTT 109

RESULT 13
US-09-815-242-5542

Sequence 5542, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5542
LENGTH: 215
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5542

Query Match 68.1%; Score 47; DB 10; Length 215;
Best Local Similarity 69.2%; Pred. No. 4.3;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ILLRGLPGSGKTT 14
Db 3 ILLMGLPGAGKGT 15

RESULT 14
US-09-815-242-12227
Sequence 12227, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12227
LENGTH: 224
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12227

Query Match 68.1%; Score 47; DB 10; Length 224;
Best Local Similarity 69.2%; Pred. No. 4.4;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ILLRGLPGSGKTT 14
Db 12 ILLMGLPGAGKGT 24

RESULT 15
US-09-815-242-12824
Sequence 12824, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12824
LENGTH: 224
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12824

Query Match 68.1%; Score 47; DB 10; Length 224;
Best Local Similarity 69.2%; Pred. No. 4.4;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ILLRGLPGSGKTT 14
Db 12 ILLMGLPGAGKGT 24

Search completed: March 10, 2003, 18:19:50

Job time : 11.7333 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2003, 17:54:59 ; Search time 11.6667 Seconds
(Without alignments)
35.308 Million cell updates/sec

Title: US-09-924-654-4_COPY_404_417
Perfect score: 69
Sequence: 1 LILLRGLPGSGKRTT 14

Scoring table:
BLOSOM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCITUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pcp.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	ID	Description
1	69	100.0	583	2	US-08-616-392C-4
2	56	81.2	394	4	US-09-195-188-1
3	50	72.5	479	4	US-09-446-504-64
4	50	72.5	479	4	US-09-712-286-64
5	50	72.5	806	1	US-07-980-528-2
6	50	72.5	1255	3	US-08-947-823-3
7	50	72.5	1257	3	US-08-947-823-5
8	47	68.1	222	4	US-09-134-001C-3289
9	47	68.1	815	4	US-09-134-001C-3614
10	46	66.7	65	4	US-09-268-364-6
11	45	65.2	223	4	US-09-574-141A-69
12	45	65.2	229	4	US-09-134-001C-4113
13	45	65.2	318	2	US-08-671-947-2
14	45	65.2	670	2	US-08-366-547-2
15	45	65.2	670	4	US-08-292-858B-2
16	44	63.8	181	1	US-08-418-444A-7
17	44	63.8	233	4	US-09-161-662-2
18	44	63.8	233	4	US-08-993-825-2
19	44	63.8	485	4	US-09-134-001C-4294
20	44	63.8	523	2	US-08-923-772-2
21	44	63.8	523	4	US-09-385-287-2
22	44	63.8	652	4	US-08-987-123-2
23	43	62.3	154	4	US-09-134-001C-3873
24	43	62.3	221	4	US-09-291-170A-8
25	43	62.3	308	4	US-09-724-884-8
26	43	62.3	308	4	US-09-149-476-607
27	43	62.3	341	4	US-09-134-001C-3650

28	43	62.3	399	4	US-09-149-476-474	Sequence 474, App
29	43	62.3	456	2	US-08-933-750C-5	Sequence 5, Appli
30	43	62.3	456	4	US-09-234-613-5	Sequence 5, Appli
31	43	62.3	456	4	US-09-268-364-21	Sequence 21, Appli
32	43	62.3	503	3	US-08-911-853-7	Sequence 7, Appli
33	43	62.3	503	4	US-09-479-409-7	Sequence 7, Appli
34	43	62.3	503	4	US-09-479-453-7	Sequence 7, Appli
35	43	62.3	806	4	US-08-999-774A-6	Sequence 6, Appli
36	43	62.3	2763	3	US-08-486-944-2	Sequence 2, Appli
37	42	60.9	28	2	US-08-724-354D-6	Sequence 6, Appli
38	42	60.9	28	3	US-09-270-984A-6	Sequence 6, Appli
39	42	60.9	239	4	US-08-858-207A-341	Sequence 341, App
40	42	60.9	264	4	US-09-035-382-4	Sequence 4, Appli
41	42	60.9	267	4	US-09-134-001C-3224	Sequence 3224, Ap
42	42	60.9	317	1	US-08-221-750A-13	Sequence 13, Appli
43	42	60.9	338	4	US-09-134-001C-4795	Sequence 4795, Ap
44	42	60.9	368	4	US-09-134-001C-5436	Sequence 5436, Ap
45	42	60.9	406	2	US-08-222-719-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-616-392C-4
Sequence 4, Application US/08616392C
Patent No. 5998165
GENERAL INFORMATION:
APPLICANT: Goold, Richard D.
APPLICANT: Akerblom, Ingrid E.
APPLICANT: Selhammer, Jeffrey
APPLICANT: Coleman, Roger
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES PANCA1A
TITLE OF INVENTION: AND PANCA1B ASSOCIATED WITH PANCREATIC CANCER
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/616,392C
FILING DATE: 15-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/581,240
FILING DATE: 29-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0052-1US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 583 amino acids
TYPE: amino acid
STRANDEDNESS: single
IMMEDIATE SOURCE:
LIBRARY: NMT
CLONE: 496071
US-08-616-392C-4
Query Match 100.0%; Score 69; DB 2; Length 583;
Best Local Similarity 100.0%; Pred. No. 0.0022;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LILLRGLPGSGKTT 14
:|||||
Db 404 LILLRGLPGSGKTT 417

RESULT 2

US-09-195-188-1
; Sequence 1, Application US/09195188
; Patent No. 6365722
; GENERAL INFORMATION:
; APPLICANT: Marcu, Kenneth B.
; TITLE OF INVENTION: Y2H4 AN IKK BINDING PROTEIN
; FILE REFERENCE: 178-264
; CURRENT APPLICATION NUMBER: US/09/195,188
; CURRENT FILING DATE: 1998-11-17
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-195-188-1

Query Match 81.2%; Score 56; DB 4; Length 394;
Best Local Similarity 69.2%; Pred. No. 0.18;
Matches 9; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LILLRGLPGSGKT 13
:|||||
Db 84 VVIMRGLPGSGKT 96

RESULT 3

US-09-446-504-64
; Sequence 64, Application US/09446504
; Patent No. 6218150
; GENERAL INFORMATION:
; APPLICANT: UEMORI, Takashi
; APPLICANT: SATO, Yoshimi
; APPLICANT: FUJITA, Tomoko
; APPLICANT: MIYAKE, Kazuo
; APPLICANT: MUKAI, Hiroyuki
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS
; FILE REFERENCE: 1422-408PCT
; CURRENT APPLICATION NUMBER: US/09/446,504
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: PCT/JP98/02845
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: JP 9-187496
; PRIOR FILING DATE: 1997-06-26
; PRIOR APPLICATION NUMBER: JP 9-320692
; PRIOR FILING DATE: 1997-11-27
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
US-09-446-504-64

Query Match 72.5%; Score 50; DB 4; Length 479;
Best Local Similarity 76.9%; Pred. No. 2;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ILLRGLPGSGKTT 14
:|||||
Db 46 LLAGPPGSGKTT 58

RESULT 4

US-09-712-266-64
; Sequence 64, Application US/09712266
; Patent No. 6331358
; GENERAL INFORMATION:
; APPLICANT: UEMORI, Takashi
; APPLICANT: SATO, Yoshimi
; APPLICANT: FUJITA, Tomoko
; APPLICANT: MIYAKE, Kazuo
; APPLICANT: MUKAI, Hiroyuki
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS
; FILE REFERENCE: 1422-408PCT
; CURRENT APPLICATION NUMBER: US/09/712,266
; CURRENT FILING DATE: 2000-11-15
; PRIOR APPLICATION NUMBER: US 09/446,504
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: PCT/JP98/02845
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: JP 9-187496
; PRIOR FILING DATE: 1997-06-26
; PRIOR APPLICATION NUMBER: JP 9-320692
; PRIOR FILING DATE: 1997-11-27
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
US-09-712-266-64

Query Match 72.5%; Score 50; DB 4; Length 479;
Best Local Similarity 76.9%; Pred. No. 2;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ILLRGLPGSGKTT 14
:|||||
Db 46 LLAGPPGSGKTT 58

RESULT 5

US-07-980-528-2
; Sequence 2, Application US/07980528
; Patent No. 5457026
; GENERAL INFORMATION:
; APPLICANT: Drevfuss, Gideon
; APPLICANT: Kiledjian, Megarditch
; APPLICANT: Portman, Douglas S.
; TITLE OF INVENTION: METHODS OF PROMOTING INTERMOLECULAR
; TITLE OF INVENTION: INTERACTIONS INVOLVING A NUCLEIC ACID AND MOLECULES
; TITLE OF INVENTION: USEFUL
; TITLE OF INVENTION: THEREFOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: NO. 5457026ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25; mb/MD
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/980,528
; FILING DATE: 19921120
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca, Mark

REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: UPN-0847
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 806 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-980-528-2

Query Match 72.5%; Score 50; DB 1; Length 806;
Best Local Similarity 57.1%; Pred No. 3.4;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 L1LRGLPGSGKTT 14
::: |||||
Db 480 VYMGIGLPGAGKTT 493

RESULT 6
US-08-947-823-3
Sequence 3, Application US/08947823
Patent No. 6114605
GENERAL INFORMATION:
APPLICANT: Williamson, Valerie M.
APPLICANT: Kaloshian, Isqouhi
APPLICANT: Yaghoobi, Jafar
APPLICANT: Bodeau, John
APPLICANT: Milligan, Stephen
TITLE OF INVENTION: Procedures and Materials for Confering
TITLE OF INVENTION: Pest Resistance in Plants
NUMBER OF SEQUENCES: 5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/947,823
FILING DATE: 09-OCT-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/18802
FILING DATE: 09-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/028,191
FILING DATE: 10-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-070210US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-947-823-3
Query Match 72.5%; Score 50; DB 3; Length 1255;

Best Local Similarity 64.3%; Pred. No. 5.5;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 L1LRGLPGSGKTT 14
: : : |||||
Db 544 VISITGMPSGKTT 557

RESULT 7
US-08-947-823-5
Sequence 5, Application US/08947823
Patent No. 6114605
GENERAL INFORMATION:
APPLICANT: Williamson, Valerie M.
APPLICANT: Kaloshian, Isqouhi
APPLICANT: Yaghoobi, Jafar
APPLICANT: Bodeau, John
APPLICANT: Milligan, Stephen
TITLE OF INVENTION: Procedures and Materials for Confering
TITLE OF INVENTION: Pest Resistance in Plants
NUMBER OF SEQUENCES: 5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/947,823
FILING DATE: 09-OCT-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/18802
FILING DATE: 09-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/028,191
FILING DATE: 10-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-070210US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1257 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-947-823-5
Query Match 72.5%; Score 50; DB 3; Length 1257;
Best Local Similarity 64.3%; Pred. No. 5.5;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 L1LRGLPGSGKTT 14
: : : |||||
Db 545 VISITGMPSGKTT 558

RESULT 8
US-09-134-001C-3289
Sequence 3289, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 3289
; LENGTH: 222
; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-3289

Query Match 68.1%; Score 47; DB 4; Length 222;

Best Local Similarity 69.2%; Pred. No. 2.6;
Matches 9; Conservative 2; Mismatches 2; Indels 2; Gaps 0;

QY 2 ILLRGLPGSGKTT 14

Db 10 IILMGLPGAGKGT 22

RESULT 9

US-09-134-001C-3614
; Sequence 3614, Application US/09134001C
; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 3614

; LENGTH: 815

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-3614

Query Match 68.1%; Score 47; DB 4; Length 815;

Best Local Similarity 64.3%; Pred. No. 11;
Matches 9; Conservative 3; Mismatches 2; Indels 2; Gaps 0;

QY 1 LILLRGLPGSGKTT 14

Db 360 IMLTGGPGTGKTT 373

RESULT 10

US-09-268-364-6
; Sequence 6, Application US/09268364A
; Patent No. 6204063

; GENERAL INFORMATION:

; APPLICANT: Allen, Steve

; APPLICANT: Lee, Jian Ming

; APPLICANT: Lightner, Jonathan

; APPLICANT: Odell, Joan

; TITLE OF INVENTION: PLANT GLYCOLYSIS AND RESPIRATION ENZYMES

; FILE REFERENCE: BB-1154

; CURRENT APPLICATION NUMBER: US/09/268,364A

; CURRENT FILING DATE: 1999-03-15

; EARLIER APPLICATION NUMBER: 60/079,387

; EARLIER FILING DATE: March 16, 1998

; NUMBER OF SEQ ID NOS: 21

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 6
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-268-364-6

Query Match 66.7%; Score 46; DB 4; Length 65;

Best Local Similarity 75.0%; Pred. No. 1;
Matches 9; Conservative 1; Mismatches 2; Indels 2; Gaps 0;

QY 3 LLRGLPGSGKTT 14

Db 35 LLHGPPGTGKTT 46

RESULT 11

US-09-574-141A-69

; Sequence 69, Application US/09574141A

; Patent No. 6395490

; GENERAL INFORMATION:

; APPLICANT: Gonsalves, Dennis

; APPLICANT: Meng, Baozhong

; TITLE OF INVENTION: RUPESTRIS STEM PITTING ASSOCIATED VIRUS

; TITLE OF INVENTION: NUCLEIC ACIDS, PROTEINS, AND THEIR USES

; FILE REFERENCE: 07678/035005

; CURRENT APPLICATION NUMBER: US/09/574,141A

; CURRENT FILING DATE: 2000-05-18

; PRIOR APPLICATION NUMBER: 60/047,147

; PRIOR FILING DATE: 1997-05-20

; PRIOR APPLICATION NUMBER: 60/069,902

; PRIOR FILING DATE: 1997-12-17

; PRIOR APPLICATION NUMBER: 09/081,320

; PRIOR FILING DATE: 1998-05-19

; NUMBER OF SEQ ID NOS: 97

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 69

; LENGTH: 223

; TYPE: PRT

; ORGANISM: Rupestris stem pitting associated virus

US-09-574-141A-69

Query Match 65.2%; Score 45; DB 4; Length 223;

Best Local Similarity 61.5%; Pred. No. 5.5;
Matches 8; Conservative 3; Mismatches 2; Indels 2; Gaps 0;

QY 2 ILLRGLPGSGKTT 14

Db 24 IVHVAVPGSGKTT 36

RESULT 12

US-09-134-001C-4113

; Sequence 4113, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC

; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 4113

; LENGTH: 229

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-4113

Query Match

65.2%; Score 45; DB 4; Length 229;

Best Local Similarity 69.2%; Pred. No. 5.7;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ILLRGLPGSGKTT 14
: : : : :
Db 42 ILLNGASGSGKTT 54

RESULT 13

US-08-671-947-2

Sequence 2, Application US/08671947
Patent No. 5827515

GENERAL INFORMATION:

APPLICANT: Shahab Reyoso, MItira

APPLICANT: Yamamoto, Takashi

APPLICANT: Cooper, Nicole H.

TITLE OF INVENTION: BACILLUS THURINGIENSIS SPOKULATION GENE

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSEE: SANDOZ AGRO, INC.

STREET: 975 California Avenue

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/671,947

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Marcus Wyner, Lynn

REGISTRATION NUMBER: 34,869

REFERENCE/DOCKET NUMBER: 133-0724

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/354-3588

TELEFAX: 415/857-1125

INFORMATION FOR SEQ. ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 318 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-671-947-2

Query Match 65.2%; Score 45; DB 2; Length 318;
Best Local Similarity 61.5%; Pred. No. 8.1;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 ILLRGLPGSGKTT 14
: : : : :
Db 92 ILLRGNPGTGKTT 104

RESULT 14

US-08-366-547-2

Sequence 2, Application US/08366547
Patent No. 5843737

GENERAL INFORMATION:

APPLICANT: Chen, Ian Bo

APPLICANT: Bao, Shideng

TITLE OF INVENTION: A NEW CANCER ASSOCIATED GENE, PROTEIN

TITLE OF INVENTION: EXPRESSED THEREFROM AND USES THEREOF

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS

ADDRESSEE: 6 CUSHMAN

STREET: 130 Water Street

CITY: Boston
STATE: MA
COUNTRY: US

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/366,547

FILING DATE: 30-DEC-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Eisenstein, Ronald I.

REGISTRATION NUMBER: 30628

REFERENCE/DOCKET NUMBER: 45072

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 523-3400

TELEFAX: (617) 523-6440

TELEX: 200281 STRE UR

INFORMATION FOR SEQ. ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 670 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-366-547-2

Query Match 65.2%; Score 45; DB 2; Length 670;
Best Local Similarity 57.1%; Pred. No. 18;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILLRGLPGSGKTT 14
: : : : :
Db 121 ILLITGPGCGKTT 134

RESULT 15

US-09-292-858B-2

Sequence 2, Application US/09292858B
Patent No. 6455681

GENERAL INFORMATION:

APPLICANT: Dean, Frank

APPLICANT: O'Donnell, Michael E.

TITLE OF INVENTION: DNA MOLECULES ENCODING SINGLE STRAND GAP RESPONSE

TITLE OF INVENTION: PROTEINS INVOLVED IN ACTIVATION OF A DNA REPAIR/CELL

FILE REFERENCE: 22221/1011

CURRENT APPLICATION NUMBER: US/09/292,858B

CURRENT FILING DATE: 1999-04-16

PRIOR APPLICATION NUMBER: 60/082,020

PRIOR FILING DATE: 1998-04-16

NUMBER OF SEQ. ID NOS: 27

SOFTWARE: Patentln Ver. 2.1

SEQ. ID NO: 2

LENGTH: 670

TYPE: PRT

ORGANISM: Homo sapiens

US-09-292-858B-2

Query Match 65.2%; Score 45; DB 4; Length 670;
Best Local Similarity 57.1%; Pred. No. 18;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 ILLRGLPGSGKTT 14
: : : : :
Db 121 ILLITGPGCGKTT 134

Search completed: March 10, 2003, 18:05:21
Job time : 12.6667 secs